Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/043696

International filing date: 23 December 2004 (23.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/587,942
Filing date: 14 July 2004 (14.07.2004)

Date of receipt at the International Bureau: 11 February 2005 (11.02.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

12498-99

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

February 03, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: ***60/587,942***
FILING DATE: ***July 14, 2004***
RELATED PCT APPLICATION NUMBER: ***PCT/US04/43696***

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR § 1.53(c).

For: **CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM**

Name of First Inventor: Emmanuel E. Baetge
Residence Address: 308 Sunset Drive, Encinitas, CA 92024

Name of Second Inventor: Kevin Allen D'Amour
Residence Address: 7235 Caminito Pantoja, San Diego, CA 92122

Name of Third Inventor: Alan D. Agulnick
Residence Address: 1259 Stratford Ct., San Marcos, CA 92069

Enclosed are:

(X) Specification in 65 pages.
(X) 45 sheets of drawings.
(X) Sequence Listing in 2 pages.
(X) The present application qualifies for small entity status under 37 CFR 1.27.
(X) A check in the amount of $80 to cover the filing fee is enclosed.
(X) A return prepaid postcard.

(X) The Commissioner is hereby authorized to charge any additional fees which may be required, now or in the future, or credit any overpayment to Account No. 11-1410.

(X) Please send correspondence to:

Jerry L. Hefner
Knobbe, Martens, Olson & Bear, LLP
2040 Main Street, 14th Floor
Irvine, CA 92614

Respectfully submitted,

Jerry L. Hefner
Registration No. 53,009
Customer No. 20,995
(619) 235-8550

Knobbe Martens Olson & Bear LLP
Intellectual Property Law

2040 Main Street
Fourteenth Floor
Irvine, CA 92614
Tel 949-760-0404
Fax 949-760-9502
www.kmob.com

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Attorney Docket No.	:	CYTHERA.045PR2
Applicant(s)	:	Baegte et al.
For	:	CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM
Attorney	:	Jerry L. Hefner
"Express Mail" Mailing Label No.	:	EV 370474765
Date of Deposit	:	July 14, 2004

I hereby certify that the accompanying Transmittal letter; Specification in 65 pages; 45 sheets of drawings; Sequence Listing in 2 pages; Check for Filing Fee; Return Prepaid Postcard are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and are addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Shani Austin

S:\DOCS\JLH\JLH-2797.DOC\071404

CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM

Field of the Invention

[0001] The present invention relates to the fields of medicine and cell biology. In particular, the present invention relates to compositions comprising CXCR4 and SOX17 expressing definitive endoderm cells and methods of making, isolating and using such cells.

Background

[0002] Human embryonic stem (ES) cells or human embryonic germ (EG) cells were first isolated in culture without fibroblast feeders in 1994 (Bongso et al., 1994) and with fibroblast feeders (Hogan, 1997; Labosky et al., 1994a; Labosky et al., 1994b). Later, Thomson, Reubinoff and Shamblott established continuous cultures of human ES and EG cells using mitotically inactivated mouse feeder layers (Reubinoff et al., 2000; Shamblott et al., 1998; Thomson et al., 1998).

[0003] Human ES and EG cells (hESCs) offer unique opportunities for investigating early stages of human development as well as for therapeutic intervention in several disease states, such as diabetes mellitus and Parkinson's disease. For example, the use of insulin-producing β-cells derived from hESCs would offer a vast improvement over current cell therapy procedures which utilize cells from donor pancreases. Currently cell therapy treatments for diabetes mellitus, which utilize cells from donor pancreases, are limited by the scarcity of high quality islet cells needed for transplant. Cell therapy for a single Type I diabetic patient requires a transplant of approximately 8×10^8 pancreatic islet cells. (Shapiro et al., 2000; Shapiro et al., 2001a; Shapiro et al., 2001b). As such, at least two healthy donor organs are required for to obtain sufficient islet cells for a successful transplant. HESCs offer a source of starting material from which to develop substantial quantities of high quality differentiated cells for human cell therapies.

[0004] Two properties that make hESCs uniquely suited to cell therapy applications are pluripotence and the ability to culture for prolonged periods without

accumulation of genetic changes. Pluripotence is defined by the ability of hESCs to differentiate to derivatives of all 3 primary germ layers (endoderm, mesoderm, ectoderm) which, in turn, form all cell types of the mature organism. Although pluripotency imparts extraordinary utility upon hESCs, this property also poses unique challenges for the study and manipulation of these cells and their derivatives. Owing to the large variety of cell types that may arise in differentiating hESC cultures, the vast majority of cell types are produced at very low efficiencies. Additionally, success in evaluating production of any given cell type depends critically on defining appropriate markers. Achieving efficient, directed differentiation is of great importance for therapeutic application of hESCs.

[0005] In order to use hESCs in cell therapy, it would be useful to overcome the foregoing problems. For example, in order to achieve the level of cellular material required for islet cell transplant therapy, it would be useful to efficiently direct hESCs toward the pancreatic islet/ β-cell lineage at the very earliest stages of differentiation.

[0006] In addition to efficient direction of the differentiation process, it is also useful to achieve isolation of intermediate cell types along the differentiation path towards the pancreatic islet/ β-cell lineage. This allows a more definitive characterization of the cell type being isolated and permits further steps in the differentiation process to be derived entirely from the appropriate lineage precursor(s).

Summary of the Invention

[0007] One embodiment of the present invention relates to novel, defined processes for the production of definitive endoderm cells in culture using pluripotent cells such as stem cells. In certain preferred embodiments, the definitive endoderm cells are derived from hESCs. Such processes provide the basis for efficient production of endodermal derived tissues such as pancreas, liver, lung, stomach, intestine and thyroid. For example, production of definitive endoderm may be the first step in differentiation of a stem cell to a functional insulin-producing β-cell. Although high efficiency of differentiation is useful for all of the many fate decisions that occur prior to reaching the pancreatic islet/ β-cell fate as shown in Figure 1, high efficiency differentiation of stem cells to definitive endoderm

cells is important because it represents an early step towards the production of functional pancreatic islet/beta cells.

[0008] Another embodiment of the present invention relates to defining the process by which definitive endoderm is first formed. Using the defined process, enriched populations of definitive endoderm cells and or tissues can be produced *in vitro* from pluripotent cells, such as stem cells. Some aspects relate to *in vitro* methodology for the production of definitive endoderm from human embryonic stem cells. For example, one method encompasses the application of culture and growth factor conditions in a defined and temporally specified fashion.

[0009] In some embodiments of the present invention, the amount or proportion of definitive endoderm cells present in a cell culture or cell population is determined. In such embodiments, the amount of transcript produced by certain genetic markers, such as SOX17, CXCR4 and other markers described herein is determined by quantitative PCR (Q-PCR). In some embodiments, immunohistochemistry to detect the proteins expressed by the above-mentioned genes. In other embodiments, Q-PCR and immunohistochemical techniques are both used to identify and determine the amount or proportion of definitive endoderm cell present in a cell culture or cell population.

[0010] In some embodiments, the definitive endoderm differentiation procedure described herein results in approximately 50-80% conversion of pluripotent cells, which include, but are not limited to, undifferentiated human stem cells, to definitive endoderm. In other embodiments, the differentiation procedures described herein result in approximately 5%, approximately 10%, approximately 20%, approximately 30%, approximately 40%, approximately 50%, approximately 60%, approximately 70%, approximately 80%, approximately 90%, or approximately 95% conversion of pluripotent cells to definitive endoderm. In other embodiments of the present invention, conversion of a pluripotent cell population, such as a stem cell population, to substantially pure definitive endoderm cell population is contemplated. In some embodiments, the definitive endoderm cells that are produced express the SOX17 gene at a level of about 2 orders of magnitude greater than non-definitive endoderm cell types.

[0011] Some embodiments of the present invention relate to compositions which comprise both stem cells and definitive endoderm cells. In some embodiments the number of stem cells is greater than the number of definitive endoderm cells, whereas in other embodiments, the number of definitive endoderm cells is greater than the number of stem cells. In some embodiments, the number of stem cells and definitive endoderm cells is approximately equivalent.

[0010] Other embodiments of the present invention include compositions comprising stem cells, definitive endoderm cells and one or more growth factors. In some embodiments, the one or more growth factors comprise the Nodal/Activin and BMP subgroups of the TGFβ superfamily of growth factors. In some embodiments, the one or more growth factors are selected from the group consisting of Nodal, Activin A, Activin B, BMP4, Wnt3a or combinations of any of these growth factors.

[0012] Another embodiment of the present invention relates to the discovery of a cell surface marker useful for the isolation and purification of definitive endoderm cells. A related embodiment is a method for using reagents which specifically bind to this marker for the enrichment, isolation, separation or purification of definitive endoderm (DE).

[0013] Still other embodiments of the present invention relate to isolated or substantially purified definitive endoderm cells. In some embodiments, the isolated or substantially purified definitive endoderm cells express the SOX17 and/or the CXRC4 marker to a greater extent than the SOX7, the AFP and/or the SPARC markers.

[0014] Other embodiments of the present inventions are described with reference to the numbered paragraphs below:

[0015] 1. A composition comprising stem cells and definitive endoderm cells, wherein at least about 5 definitive endoderm cells are present for about every 95 stem cells present in said composition.

[0016] 2. The composition of Paragraph 1, wherein at least about 10 definitive endoderm cells are present for about every 90 stem cells present in said composition.

[0017] 3. The composition of Paragraph 1, wherein at least about 20 definitive endoderm cells are present for about every 80 stem cells present in said composition.

[0018] 4. The composition of Paragraph 1, wherein at least about 30 definitive endoderm cells are present for about every 70 stem cells present in said composition.

[0019] 5. The composition of Paragraph 1, wherein at least about 40 definitive endoderm cells are present for about every 60 stem cells present in said composition.

[0020] 6. The composition of Paragraph 1, wherein at least about 50 definitive endoderm cells are present for about every 50 stem cells present in said composition.

[0021] 7. The composition of Paragraph 1, wherein at least about 60 definitive endoderm cells are present for about every 40 stem cells present in said composition.

[0022] 8. The composition of Paragraph 1, wherein at least about 70 definitive endoderm cells are present for about every 30 stem cells present in said composition.

[0023] 9. The composition of Paragraph 1, wherein at least about 80 definitive endoderm cells are present for about every 20 stem cells present in said composition.

[0024] 10. The composition of Paragraph 1, wherein at least about 90 definitive endoderm cells are present for about every 10 stem cells present in said composition.

[0025] 11. The composition of Paragraph 1, wherein at least about 95 definitive endoderm cells are present for about every 5 stem cells present in said composition.

[0026] 12. The composition of Paragraph 1, wherein said stem cells are embryonic stem cells.

[0027] 13. The composition of Paragraph 12, wherein said embryonic stem cells are derived from the inner cell mass of an embryo.

14. The composition of Paragraph 12, wherein said embryonic stem cells are derived from the gonadal ridges of an embryo.

[0028] 15. The composition of Paragraph 1, wherein said stem cells are of human origin.

[0029] 16. The composition of Paragraph 1, wherein said definitive endoderm cells are of human origin.

[0030] 17. The composition of Paragraph 1, wherein said definitive endoderm cells express the SOX17 gene.

[0031] 18. The composition of Paragraph 17, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the SOX7 gene.

[0032] 19. The composition of Paragraph 18, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the AFP gene.

[0033] 20. The composition of Paragraph 19, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the SPARC gene.

[0034] 21. The composition of Paragraph 20, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the Thrombomodulin gene.

[0035] 22. The composition of Paragraph 21, wherein said definitive endoderm cells express the MIXL1 gene.

[0036] 23. The composition of Paragraph 21, wherein said definitive endoderm cells express the GATA4 gene.

[0037] 24. The composition of Paragraph 21, wherein said definitive endoderm cells express the HNF3b gene.

[0038] 25. The composition of Paragraph 1 further comprising a growth factor of the Nodal/Activin subgroup of the TGFβ superfamily.

[0039] 26. The composition of Paragraph 1 further comprising a growth factor of the BMP subgroup of the TGFβ superfamily.

[0040] 27. The composition of Paragraph 1 further comprising a growth factor selected from the group consisting of Nodal, Activin A, Activin B, BMP4, Wnt3a and combinations thereof.

[0041] 28. The composition of Paragraph 1 further comprising Nodal, Activin A, Activin B and BMP4.

[0042] 29. A substantially purified definitive endoderm cell.

[0043] 30. The definitive endoderm cell of Paragraph 29, wherein said cell is derived from an embryonic stem cell.

[0044] 31. The definitive endoderm cell of Paragraph 29, wherein said cell expresses the SOX17 gene.

[0045] 32. The definitive endoderm cell of Paragraph 31, wherein the expression of the SOX17 gene is greater than the expression of the SOX7 gene.

[0046] 33. The definitive endoderm cell of Paragraph 32, wherein the expression of the SOX17 gene is greater than the expression of the AFP gene.

[0047] 34. The definitive endoderm cell of Paragraph 33 wherein the expression of the SOX17 gene is greater than the expression of the SPARC gene.

[0048] 35. The definitive endoderm cell of Paragraph 34, wherein the expression of the SOX17 gene is greater than the expression of the Thrombomodulin gene.

[0049] 36. The definitive endoderm cell of Paragraph 35, wherein said cell expresses the MIXL1 gene.

[0050] 37. The definitive endoderm cell of Paragraph 35, wherein said cell expresses the GATA4 gene.

[0051] 38. The definitive endoderm cell of Paragraph 35, wherein said cell expresses the HNF3b gene.

[0052] 39. The definitive endoderm cell of Paragraph 29, wherein said cell is of human origin.

[0053] 40. A method for producing definitive endoderm cells, said method comprising obtaining a stem cell culture; and providing to said culture a growth factor of the TGFβ superfamily in an amount sufficient to promote differentiation of at least a portion of said stem cell culture to definitive endoderm cells.

[0054] 41. The method of Paragraph 40, wherein at least about 5 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0055] 42. The method of Paragraph 40, wherein at least about 10 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0056] 43. The method of Paragraph 40, wherein at least about 20 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0057] 44. The method of Paragraph 40, wherein at least about 30 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0058] 45. The method of Paragraph 40, wherein at least about 40 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0059] 46. The method of Paragraph 40, wherein at least about 50 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0060] 47. The method of Paragraph 40, wherein at least about 60 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0061] 48. The method of Paragraph 40, wherein at least about 70 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0062] 49. The method of Paragraph 40, wherein at least about 80 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0063] 50. The method of Paragraph 40, wherein at least about 90 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0064] 51. The method of Paragraph 40, wherein at least about 95 percent of the stem cells in said culture differentiates into said definitive endoderm cells.

[0065] 52. The method of Paragraph 40, wherein the stem cells in said culture are embryonic stem cells.

[0066] 53. The method of Paragraph 52, wherein said embryonic stem cells are derived from the inner cell mass of an embryo.

[0067] 54. The method of Paragraph 52, wherein said embryonic stem cells are derived from the gonadal ridges of an embryo.

[0068] 55. The method of Paragraph 40, wherein said stem cells in said culture are of human origin.

[0069] 56. The method of Paragraph 40, wherein said definitive endoderm cells are of human origin.

[0070] 57. The method of Paragraph 40, wherein said definitive endoderm cells express the SOX17 gene.

[0071] 58. The method of Paragraph 57, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the SOX7 gene.

[0072] 59. The method of Paragraph 58, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the AFP gene.

[0073] 60. The method of Paragraph 59, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the SPARC gene.

[0074] 61. The method of Paragraph 60, wherein the expression of the SOX17 gene in said definitive endoderm cells is greater than the expression of the Thrombomodulin gene.

[0075] 62. The method of Paragraph 61, wherein said definitive endoderm cells express the MIXL1 gene.

[0076] 63. The method of Paragraph 61, wherein said definitive endoderm cells express the GATA4 gene.

[0077] 64. The method of Paragraph 61, wherein said definitive endoderm cells express the HNF3b gene.

[0078] 65. The method of Paragraph 40, wherein said growth factor is of the Nodal/Activin subgroup of the TGFβ superfamily.

[0079] 66. The method of Paragraph 40, wherein said growth factor is of the BMP subgroup of the TGFβ superfamily.

[0080] 67. The method of Paragraph 40, wherein said growth factor is selected from the group consisting of Nodal, Activin A, Activin B, BMP4 and combinations thereof.

[0081] 68. The method of Paragraph 67 further comprising the growth factor Wnt3a.

[0082] 69. The method of Paragraph 40, wherein a plurality of growth factors of the TGFβ superfamily is provided.

[0083] 70. The method of Paragraph 69, wherein the plurality of growth factors comprises Nodal, Activin A, Activin B and BMP4.

[0084] 71. The method of Paragraph 40, wherein said growth factor is BMP4.

[0085] 72. The method of Paragraph 71 further comprising the step of removing BMP4 within approximately 4 days after its addition.

[0086] 73. The method of Paragraph 40 wherein said growth factor is provided in a concentration of at least about 10 ng/ml.

[0087] 74. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 25 ng/ml.

[0088] 75. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 50 ng/ml.

[0089] 76. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 75 ng/ml.

[0090] 77. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 100 ng/ml.

[0091] 78. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 200 ng/ml.

[0092] 79. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 300 ng/ml.

[0093] 80. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 400 ng/ml.

[0094] 81. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 500 ng/ml.

[0095] 82. The method of Paragraph 40, wherein said growth factor is provided in a concentration of at least about 1000 ng/ml.

[0096] 83. A definitive endoderm cell produced by the method of Paragraph 40.

[0097] 84. An antibody which binds to SOX17.

[0098] 85. The antibody of Paragraph 84, wherein said SOX17 is human SOX17.

[0099] 86. The antibody of Paragraph 84, wherein said antibody is a polyclonal.

[0100] 87. The antibody of Paragraph 84, wherein said antibody is a monoclonal antibody.

[0101] 88. A method of increasing the expression of the SOX17 gene product in a stem cell comprising contacting said stem cell with a growth factor of the TGFβ superfamily in an amount sufficient to increase expression of the SOX17 gene product.

[0102] 89. The method of Paragraph 88, wherein the stem cell is an embryonic stem cell.

[0103] 90. The method of Paragraph 89, wherein said embryonic stem cell is derived from the inner cell mass of an embryo.

[0104] 91. The method of Paragraph 89, wherein said embryonic stem cell is derived from the gonadal ridges of an embryo.

[0105] 92. The method of Paragraph 88, wherein said stem cell is of human origin.

[0106] 93. The method of Paragraph 88, wherein said growth factor is selected from the group consisting of Nodal, Activin A, Activin B, BMP4 and combinations thereof.

[0107] 94. The method of Paragraph 93 further comprising the growth factor Wnt3a.

[0108] 95. The method of Paragraph 88, wherein a plurality of growth factors of the TGFβ superfamily is provided.

[0109] 96. The method of Paragraph 95, wherein the plurality of growth factors comprises Nodal, Activin A, Activin B and BMP4.

[0110] 97. The method of Paragraph 88, wherein said growth factor is BMP4.

[0111] 98. The method of Paragraph 97 further comprising the step of removing BMP4 within approximately 4 days after its addition.

[0112] 99. A mammalian cell culture comprising endodermal cells, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 10% of said endodermal cells.

[0113] 100. The cell culture of Paragraph 99, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 20% of said endodermal cells.

[0114] 101. The cell culture of Paragraph 99, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 40% of said endodermal cells.

[0115] 102. The cell culture of Paragraph 99, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a

marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 50% of said endodermal cells.

[0116] 103. The cell culture of Paragraph 99, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 70% of said endodermal cells.

[0117] 104. The cell culture of Paragraph 99, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 90% of said endodermal cells.

[0118] 105. A mammalian cell culture comprising endodermal cells, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 10% of said endodermal cells.

[0119] 106. The cell culture of Paragraph 105, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 20% of said endodermal cells.

[0120] 107. The cell culture of Paragraph 105, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 40% of said endodermal cells.

[0121] 108. The cell culture of Paragraph 105, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 50% of said endodermal cells.

[0122] 109. The cell culture of Paragraph 105, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 70% of said endodermal cells.

[0123] 110. The cell culture of Paragraph 105, wherein the expression of SOX17 and CXCR4 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7 in at least about 90% of said endodermal cells.

[0124] 111. A mammalian cell culture comprising endodermal cells, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 10% of said endodermal cells.

[0125] 112. The cell culture of Paragraph 111, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 20% of said endodermal cells.

[0126] 113. The cell culture of Paragraph 111, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 40% of said endodermal cells.

[0127] 114. The cell culture of Paragraph 111, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 60% of said endodermal cells.

[0128] 115. The cell culture of Paragraph 111, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 80% of said endodermal cells.

[0129] 116. The cell culture of Paragraph 111, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 90% of said endodermal cells.

[0130] 117. A mammalian cell culture comprising endodermal cells, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 10% of said endodermal cells.

[0131] 118. The cell culture of Paragraph 117, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 20% of said endodermal cells.

[0132] 119. The cell culture of Paragraph 117, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 40% of said endodermal cells.

[0133] 120. The cell culture of Paragraph 117, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 60% of said endodermal cells.

[0134] 121. The cell culture of Paragraph 117, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 80% of said endodermal cells.

[0135] 122. The cell culture of Paragraph 117, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers in at least about 90% of said endodermal cells.

[0136] 123. A method of producing a cell population enriched in definitive endoderm cells, said method comprising the steps of:

differentiating cells in a population comprising pluripotent cells so as to produce definitive endoderm cells;

providing to said cell population a reagent which binds to a marker expressed in a definitive endoderm cell but which is not substantially expressed in other cell types present in said cell population; and

separating said definitive endoderm cells bound to said reagent from said other cell types present in said cell population, thereby producing a cell population enriched in definitive endoderm cells.

[0137] 124. The method of Paragraph 123, wherein said pluripotent cells are stem cells.

[0138] 125. The method of Paragraph 124, wherein said stem cells are human embryonic stem cells.

[0139] 126. The method of Paragraph 123, wherein said step of differentiating said cell population comprises providing to said cell population a growth factor of the TGFβ superfamily in an amount sufficient to promote differentiation of at least a portion of said cell population to definitive endoderm cells.

[0140] 127. The method of Paragraph 126, wherein said growth factor is selected from the group consisting of Nodal, Activin A, Activin B, BMP4 and combinations thereof.

[0141] 128. The method of Paragraph 126, wherein the growth factor is Wnt3a.

[0142] 129. The method of Paragraph 126, wherein a plurality of growth factors of the TGFβ superfamily is provided.

[0143] 130. The method of Paragraph 129, wherein the plurality of growth factors comprises Nodal, Activin A, Activin B and BMP4.

[0144] 131. The method of Paragraph 123, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in said definitive endoderm cells.

[0145] 132. The method of Paragraph 123, wherein said marker is SOX17.

[0146] 133. The method of Paragraph 123, wherein said marker is CXCR4.

[0147] 134. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 50% definitive endoderm cells

[0148] 135. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 60% definitive endoderm cells

[0149] 136. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 70% definitive endoderm cells

[0150] 137. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 80% definitive endoderm cells

[0151] 138. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 90% definitive endoderm cells

[0152] 139. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 95% definitive endoderm cells

[0153] 140. The method of Paragraph 123, wherein said cell population enriched in definitive endoderm cells comprises at least about 98% definitive endoderm cells

[0154] 141. A substantially purified mammalian cell composition comprising endodermal cells, wherein expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7.

[0155] 142. A substantially purified mammalian cell composition comprising endodermal cells, wherein the expression of SOX17 and CXCR4 is greater than the expression of the alpha-fetoprotein (AFP) and Thrombomodulin (TM) markers.

[0156] 143. A substantially purified mammalian cell composition comprising cells, in which expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of a marker selected from the group consisting of alpha-fetoprotein (AFP), Thrombomodulin (TM) and SOX7, wherein said cells have been obtained by inducing pluripotent cells to differentiate.

[0157] 144. The cell cultures of any one of Paragraphs 99-110, wherein said SOX17, CXCR4, AFP, TM and SOX7 expression is determined by quantitative PCR (Q-PCR).

[0158] 145. The cell cultures of any one of Paragraphs 99-110, wherein said SOX17, CXCR4, AFP, TM and SOX7 expression is determined by immunocytochemistry.

[0159] 146. The cell cultures of any one of Paragraphs 111-122, wherein said SOX17, CXCR4, AFP and TM expression is determined by quantitative PCR (Q-PCR).

[0160] 147. The cell cultures of any one of Paragraphs 111-122, wherein said SOX17, CXCR4, AFP and TM expression is determined by immunocytochemistry.

[0161] 148. The method of Paragraph 131, wherein said SOX17, CXCR4, AFP and TM expression is determined by quantitative PCR (Q-PCR).

[0162] 149. The method of Paragraph 131, wherein said SOX17, CXCR4, AFP and TM expression is determined by immunocytochemistry.

[0163] 150. The substantially purified mammalian cell compositions of any one of Paragraphs 131-143, wherein said SOX17, CXCR4, AFP and TM expression is determined by quantitative PCR (Q-PCR).

[0164] 151. The substantially purified mammalian cell compositions of any one of Paragraphs 131-143, wherein said SOX17, CXCR4, AFP and TM expression is determined by immunocytochemistry.

[0165] It will be appreciated that the methods and compositions described above relate cells cultured *in vitro*. However, the above-described *in vitro* differentiated cell compositions may be used for *in vivo* applications.

[0166] Additional embodiments of the present inventions may also be found in United States Provisional Patent Application No. 60/532,004, entitled DEFINITIVE ENDODERM, filed December 23, 2003; United States Provisional Patent Application No. 60/566,293, entitled PDX1 EXPRESSING ENDODERM, filed April 27, 2004 and the United States provisional patent application corresponding to attorney docket number CYTHERA.045PR, entitled CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM, filed July 9, 2004, the disclosures of which are incorporated herein by reference in their entireties.

Brief Description of the Drawings

[0167] Figure 1 is a schematic of a proposed differentiation pathway for the production of beta-cells from hESCs. The first step in the pathway commits the ES cell to the definitive endoderm lineage and represents an essential step prior to further differentiation events to pancreatic endoderm, endocrine endoderm, or islet/beta-cell. Some factors useful for mediating this transition are highlighted in red. Relevant markers for defining the definitive endoderm target cell are listed in blue.

[0168] Figure 2A is a micrograph showing characteristic morphology of an undifferentiated hESC colony.

[0169] Figure 2B is an inset from Figure 2A showing the characteristic morphology (round with prominent nucleoli) of individual hESCs at 200X magnification.

[0170] Figure 3 is an image of the karyotype of hESCyT-25 at passage 10 and 28. hESCyT-25 displays a normal female karyotype that remains stable over time through passages 10 and 28.

[0171] Figures 4A-C are micrographs showing that hESCyT-25 cells display immunoreactivity for the characteristic markers of undifferentiated hESCs. (A) TRA-1-60, (B) OCT4, (C) SSEA4.

[0172] Figures 5A-D are micrographs which demonstrate that undifferentiated hESCyT-25 cells are recognizable by bright field morphology of colonies (A) and individual cells (C). It is also demonstrated that these cells display the hESC characteristic of robust alkaline phosphatase activity both in colonies (B) and individual cells (D).

[0173] Figure 6 is a micrograph which demonstrates that hESCyT-25 cells can be differentiated as floating embryoid bodies when cultured on low adherent substrata.

[0174] Figure 7 is a micrograph which shows the differentiation of hESCyT-25 to neurons (ectoderm germ layer) as indicated by immunoreactivity to βIII-tubulin.

[0175] Figure 8A is a line graph showing differentiation of hESCyT-25 to endoderm germ layer is indicated by the robust expression of SOX7 and AFP.

[0176] Figure 8B is a micrograph showing cell differentiated in monolayer which displays regions of strong immunoreactivity for AFP.

[0177] Figure 9 is a bar chart showing rapid increase in expression of brachyury with time under differentiating suspension culture conditions. Increased expression of brachyury is indicative of differentiation of hESCyT-25 to cells of mesoderm lineage.

[0178] Figure 10 is a diagram of the human SOX17 cDNA which displays the positions of conserved motifs and highlights the region used for the immunization procedure by GENOVAC.

[0179] Figure 11 is a Western blot probed with the rat anti-SOX17 antibody. This blot demonstrates the specificity of this antibody for human SOX17 protein over-expressed in fibroblasts (lane 1) and a lack of immunoreactivity with EGFP (lane 2) or the most closely related SOX family member, SOX7 (lane 3).

[0180] Figures 12A-F are micrographs which demonstrate SOX17 antibody specificity by immunocytochemistry. Panels A-C show fibroblasts which were transfected

with human SOX17 and EGFP cDNA whereas panels D-F show fibroblasts transfected with EGFP alone. SOX17 immunoreactive cells (red) are observed only in SOX17 transfected cultures (A and C but not D and F) and reactivity correlates with the EGFP signal (green – B and E) indicating that those cells were transfected (C).

[0181] Figure 13 is a relational dendrogram illustrating that SOX17 is most closely related to SOX7 and somewhat less to SOX18. The SOX17 proteins are more closely related among species homologs than to other members of the SOX group F subfamily within the same species.

[0182] Figures 14A-B are micrographs showing a cluster of $SOX17^+$ (red) cells that display a significant number of AFP^+ (green) co-labeled cells (A). This is in striking contrast to other $SOX17^+$ clusters (B) where little or no AFP^+ cells are observed.

[0183] Figures 15A-C are micrographs showing parietal endoderm and SOX17. Panel A shows immunocytochemistry (green fluorescence) for human Thrombomodulin (TM) protein located on the cell surface of parietal endoderm cells in randomly differentiated cultures of hES cells. Panel B is the identical field shown in A that is double-labeled for TM (green) and SOX17 (red). Panel C is the phase contrast image of the same field with DAPI labeled nuclei. Note the complete correlation of DAPI labeled nuclei and SOX17 labeling.

[0184] Figures 16A-B are bar charts showing SOX17 gene expression by quantitative PCR (Q-PCR) and anti-SOX17 positive cells by SOX17-specific antibody. Panel A shows that Activin A increases SOX17 gene expression while retinoic acid (RA) strongly suppresses SOX17 expression relative to the undifferentiated control media (SR20). Panel B shows the identical pattern as well as a similar magnitude of these changes is reflected in $SOX17^+$ cell number, indicating that Q-PCR measurement of SOX17 gene expression is very reflective of changes at the single cell level.

[0185] Figure 17A is a bar charts which shows that a culture of differentiating hESCs in the presence of Activin A maintains a low level of AFP gene expression while cells allowed to randomly differentiate in 10% fetal bovine serum (FBS) exhibit a strong upregulation of AFP. The difference in expression levels is approximately 7-fold.

[0186] Figure 17B is an image of two micrographs showing that the suppression of AFP expression by Activin A is also evident at the single cell level as indicated by the very

rare and small clusters of AFP^+ cells observed in Activin A treatment conditions (bottom) relative to 10% FBS alone (top).

[0187] Figure 18 is a comparative image showing the quantitation of the AFP^+ cell number using flow cytometry. This figure demonstrates that the magnitude of change in AFP gene expression (Figure 17A) in the presence (right panel) and absence (left panel) of Activin A exactly corresponds to the number of AFP^+ cells, further supporting the utility of Q-PCR analyses to indicate changes occurring at the individual cell level.

[0188] Figures 19A-F are micrographs which show that exposure of hESCs to nodal, Activin A and Activin B (NAA) yields a striking increase in the number of $SOX17^+$ cells over the period of 5 days (A-C). By comparing to the relative abundance of $SOX17^+$ cells to the total number of cells present in each field, as indicated by blue DAPI stained nuclei (D-F), it can be seen that approximately 30-50% of all cells are immunoreactive for SOX17 after five days treatment with NAA.

[0189] Figure 20 is an image displaying graphs that are representative examples of the 3-point standard curve and melt curve analysis using the Rotor Gene 3000 instrument for Q-PCR. Prerequisites for acceptable Q-PCR performance are high correlation coefficients (≥ 0.98), PCR efficiency values near 100%, and the presence of a single PCR product with no amplification from samples that did not receive reverse transcriptase.

[0190] Figure 21 is a bar chart which demonstrates that Activin A (0, 10, 30 or 100 ng/mL) dose-dependently increases SOX17 gene expression in differentiating hESCs. Increased expression is already robust after 3 days of treatment on adherent cultures and continues through subsequent 1, 3 and 5 days of suspension culture as well.

[0191] Figures 22A-C are bar charts which demonstrate the effect of Activin A on the expression of MIXL1 (panel A), GATA4 (panel B) and HNF3b (panel C). Activin A dose-dependent increases are also observed for three other markers of definitive endoderm; MIXL1, GATA4 and HNF3b. The magnitudes of increased expression in response to activin dose are strikingly similar to those observed for SOX17, strongly indicating that Activin A is specifying a population of cells that co-express all four genes ($SOX17^+$, $MIXL1^+$, $GATA4^+$ and $HNF3b^+$).

[0192] Figures 23A-C are bar charts which demonstrate the effect of Activin A on the expression of AFP (panel A), SOX7 (panel B) and SPARC (panel C). There is an Activin A dose-dependent decrease in expression of the visceral endoderm marker AFP. Markers of primitive endoderm (SOX7) and parietal endoderm (SPARC) remain either unchanged or exhibit suppression at some time points indicating that Activin A does not act to specify these extra-embryonic endoderm cell types. This further supports the fact that the increased expression of SOX17, MIXL1, GATA4, and HNF3b are due to an increase in the number of definitive endoderm cells in response to Activin A.

[0193] Figures 24A-B are bar charts showing the effect of Activin A on ZIC1 (panel A) and Brachyury expression (panel B) Consistent expression of the neural marker ZIC1 demonstrates that there is no dose-dependent effect of Activin A on neural differentiation. There is a notable suppression of mesoderm differentiation mediated by 100 ng/mL of Activin A treatment as indicated by the decreased expression of brachyury. This is likely the result of the increased specification of definitive endoderm from the mesendoderm precursors. Lower levels of Activin A treatment (10 and 30 ng/mL) maintain the expression of brachyury at later time points of differentiation relative to untreated control cultures.

[0194] Figures 25A-C are micrographs showing (A) hESCs at the periphery of the colony exhibit signs of differentiation indicated by their elongated morphologies (arrows). (B, C) These same cells also appear to have increased immunoreactivity for OCT4.

[0195] Figures 26A-F are micrographs showing that undifferentiated hESCs expressing OCT4 do not express SOX17. In hESC colonies of primarily undifferentiated phenotype (A, C, E), $SOX17^{+}$ cells occur only at the periphery where differentiated cells begin to appear and none co-label with OCT4 which is found throughout the remainder of the colony (C). In colonies where differentiation is widespread (B, D, F) clusters $SOX17^{+}$ cells are present scattered throughout the interior of the colony as well as at the edges and boundaries between colonies (D). However, even in colonies of such heterogenous state, OCT4 and SOX17 immunoreactivity is not found in the same cells although they occur highly intermingled with each other. (OCT4 – red, SOX17 – green, DAPI – blue).

[0196] Figures 27A-B are micrographs showing decreased parietal endoderm differentiation in response to treatment with activins. Regions of TM^{hi} parietal endoderm are

found through the culture (A) when differentiated in serum alone, while differentiation to TM^+ cells is scarce when activins are included (B) and overall intensity of TM immunoreactivity is lower.

[0197] Figure 28 is a micrograph showing the appearance of definitive endoderm and visceral endoderm in vitro from hESCs. The regions of visceral endoderm are identified by $AFP^{hi}/SOX17^{lo/-}$ while definitive endoderm displays the complete opposite profile, $SOX17^{hi}/AFP^{lo/-}$. It is possible that a small proportion of these $SOX17^{hi}/AFP^{lo/-}$cells may be parietal endoderm (see text). This field was selectively chosen due to the proximity of these two regions to each other. However, there are numerous times when $SOX17^{hi}/AFP^{lo/-}$ regions are observed in absolute isolation from any regions of AFP^{hi} cells, suggesting the separate origination of the definitive endoderm cells from visceral endoderm cells.

[0198] Figure 29 is diagram depicting the TGFβ family of ligands and receptors. Factors activating AR Smads and BR Smads are useful in the production of definitive endoderm from human embryonic stem cells (see, *J Cell Physiol.* **187**:265-76).

[0199] Figure 30 is a bar chart showing the induction of SOX17 expression over time as a result of treatment with individual and combinations of TGFβ factors.

[0200] Figure 31 is a bar chart showing the increase in $SOX17^+$ cell number with time as a result of treatment with combinations of TGFβ factors.

[0201] Figure 32 is a bar chart showing induction of SOX17 expression over time as a result of treatment with combinations of TGFβ factors.

[0202] Figure 33 is a bar chart showing that Activin A induces a dose-dependent increase in $SOX17^+$ cell number.

[0203] Figure 34 is a bar chart showing that addition of Wnt3a to Activin A and Activin B treated cultures increases SOX17 expression above the levels induced by Activin A and Activin B alone.

[0204] Figures 35A-C are bar charts showing differentiation to definitive endoderm is enhanced in low FBS conditions. Treatment of hESCs with activins A and B in media containing 2% FBS (2AA) yields a 2-3 times greater level of SOX17 expression as compared to the same treatment in 10% FBS media (10AA) (panel A). Induction of the definitive endoderm marker MIXL1 (panel B) is also affected in the same way and the

suppression of AFP (visceral endoderm) (panel C) is greater in 2% FBS than in 10% FBS conditions.

[0205] Figures 36A-D are micrographs which show $SOX17^+$ cells are dividing in culture. SOX17 immunoreactive cells (green) are present at the differentiating edge of an hESC colony (C, D) and are labeled with proliferating cell nuclear antigen (PCNA) (red in panel B) yet are not co-labeled with OCT4 (red in panel C). In addition, clear mitotic figures can be seen by DAPI labeling of nuclei in both $SOX17^+$ cells (arrows) as well as $OCT4^+$, undifferentiated hESCs (arrowheads) (D).

[0206] Figures 37A-D are micrographs which show hESC marker expression in response to treatment with Activin A and Activin B. hESCs were treated for four consecutive days with Activin A and Activin B and triple labeled with SOX17, AFP and TM antibodies. Panel A - SOX17; Panel B - AFP; Panel C - TM; and Panel D - Phase/DAPI. Notice the numerous SOX17 positive cells (A) associated with the complete absence of AFP (B) and TM (C) immunoreactivity.

[0207] Figure 38 is a bar chart showing the relative expression level of CXCR4 in differentiating hESCs under various media conditions.

[0208] Figures 39A-D are bar charts that show how a panel of definitive endoderm markers share a very similar pattern of expression to CXCR4 across the same differentiation treatments displayed in Figure 38.

[0209] Figures 40A-E are bar charts showing how markers for mesoderm (BRACHYURY, MOX1), ectoderm (SOX1, ZIC1) and visceral endoderm (SOX7) exhibit an inverse relationship to CXCR4 expression across the same treatments displayed in Figure 38.

[0210] Figures 41A-C are micrographs that show the relative difference in SOX17 immunoreactive cells across three of the media conditions displayed in Figures 38-40.

[0211] Figures 42A-C are flow cytometry dot plots that demonstrate the increase in $CXCR4^+$ cell number with increasing concentration of activin A added to the differentiation media.

[0212] Figure 43 is a bar chart showing gene expression from $CXCR4^+$ and $CXCR4^-$ cells isolated using fluorescence-activated cell sorting (FACS) as well as gene

expression in the parent populations. This demonstrates that the $CXCR4^+$ cells contain essentially all the CXCR4 gene expression present in each parent population and the $CXCR4^-$ populations contain very little or no CXCR4 gene expression.

[0213] Figures 44A-D are bar charts that show the $CXCR4^+$ cells isolated from the high dose activin A treatment (A100-CX+) are even further enriched for definitive endoderm markers than the parent population (A100).

[0214] Figures 45A-D are bar charts that demonstrate the depletion of mesoderm (BRACHYURY, MOX1), ectoderm (ZIC1) and visceral endoderm (SOX7) gene expression in the CXCR4+ cells isolated from the high dose activin A treatment which is already suppressed in expression of these non-definitive endoderm markers.

Detailed Description

[0215] In accordance with some embodiments of the present invention, methods of producing definitive endoderm from pluripotent cells, such as pluripotent stem cells, are disclosed. Stem cells used in these methods can include, but are not limited to, embryonic stem cells. Embryonic stem cells can be derived from the embryonic inner cell mass or from the embryonic gonadal ridges. Embryonic stem cells can originate from a variety of animal species including, but not limited to, various mammalian species including humans.

[0216] In some embodiments of the present invention, one or more growth factors are used in the differentiation process from stem cell to definitive endoderm cell. Such factors can include growth factors from the BMP subgroup of the TGFβ superfamily. For example, such factors include, but are not limited to Nodal, Activin A, Activin B, BMP4 and combinations thereof. The use of the growth factor Wnt3a is also contemplated.

[0217] Other aspects of the invention disclosed herein relate to compositions which comprise both stem cells and definitive endoderm cells. In some embodiments, such compositions also include one or more growth factors. In other embodiments, compositions described herein are enriched for definitive endoderm cells or comprise substantially purified definitive endoderm cells.

[0218] In some embodiments of the present invention, SOX17 antibodies, CXCR4 antibodies, SDF-1 ligands or other ligands for CXCR4 can be used to obtain

definitive endoderm cells in an enriched, isolated or substantially purified form. For example, a SOX17 antibody, a CXCR4 antibody, an SDF-1 ligand or another ligand for CXCR4 can be used as a reagent in a method, such as affinity-based separation or magnetic-based separation, to enrich, isolate or substantially purify preparations of definitive endoderm cells which bind to the reagent.

[0219] The compositions and methods described herein have several useful features. For example, the compositions and methods described herein are useful for modeling the early stages of human development. Furthermore, the compositions and methods described herein can also serve for therapeutic intervention in disease states, such as diabetes mellitus. For example, since definitive endoderm serves as the source for only a limited number tissues, it can be used in the development of pure tissue or cell types.

[0220] A crucial stage in early human development termed gastrulation occurs 2-3 weeks after fertilization. Gastrulation is extremely significant because it is at this time that the three primary germ layers are first specified and organized (Lu et al., 2001; Schoenwolf and Smith, 2000). The ectoderm is responsible for the eventual formation of the outer coverings of the body and the entire nervous system whereas the heart, blood, bone, skeletal muscle and other connective tissues are derived from the mesoderm. Definitive endoderm is defined as the germ layer that is responsible for formation of the entire gut tube which includes the esophagus, stomach and small and large intestines, and the organs which derive from the gut tube such as the lungs, liver, thymus, parathyroid and thyroid glands, gall bladder and pancreas (Grapin-Botton and Melton, 2000; Kimelman and Griffin, 2000; Tremblay et al., 2000; Wells and Melton, 1999; Wells and Melton, 2000). A very important distinction should be made between the definitive endoderm and the completely separate lineage of cells termed primitive endoderm. The primitive endoderm is primarily responsible for formation of extra-embryonic tissues, mainly the parietal and visceral endoderm portions of the placental yolk sac and the extracellular matrix material of Reichert's membrane.

[0221] During gastrulation, the process of definitive endoderm formation begins with a cellular migration event in which mesendoderm cells (cells competent to form mesoderm or endoderm) migrate through a structure called the primitive streak. Definitive endoderm is derived from cells, which migrate through the anterior portion of the streak and

through the node (a specialized structure at the anterior-most region of the streak). As migration occurs, definitive endoderm populates first the most anterior gut tube and culminates with the formation of the posterior end of the gut tube.

[0222] *In vivo* analyses of the formation of definitive endoderm, such as the studies in Zebrafish and *Xenopus* by Conlon et al., 1994; Feldman et al., 1998; Zhou et al., 1993; Aoki et al., 2002; Dougan et al., 2003; Tremblay et al., 2000; Vincent et al., 2003; Alexander et al., 1999; Alexander and Stainier, 1999; Kikuchi et al., 2001; Hudson et al., 1997 and in mouse by Kanai-Azuma et al., 2002 lay a foundation for how one might attempt to approach the development of a specific germ layer cell type in the culture dish using human embryonic stem cells. There are two aspects associated with in vitro ESC culture that pose major obstacles in the attempt to recapitulate development in the culture dish. First, organized germ layer or organ structures are not produced. The majority of germ layer and organ specific genetic markers will be expressed in a heterogeneous fashion in the differentiating hESC culture system. Therefore it is difficult to evaluate formation of a specific tissue or cell type due to this lack of organ specific boundaries. Almost all genes expressed in one cell type within a particular germ layer or tissue type are expressed in other cells of different germ layer or tissue types as well. Without specific boundaries there is considerably less means to assign gene expression specificity with a small sample of 1-3 genes. Therefore one must examine considerably more genes, some of which should be present as well as some that should not be expressed in the particular cell type of the organ or tissue of interest. Second, the timing of gene expression patterns is crucial to movement down a specific developmental pathway.

[0223] To further complicate matters, it should be noted that stem cell differentiation *in vitro* is rather asynchronous, likely considerably more so than *in vivo*. As such, one group of cells may be expressing genes associated with gastrulation, while another group maybe starting final differentiation. Furthermore, manipulation of hESC monolayers or embryoid bodies (EBs) with or without exogenous factor application may result in profound differences with respect to overall gene expression pattern and state of differentiation. For these reasons, the application of exogenous factors must be timed

according to gene expression patterns within a heterogeneous cell mixture in order to efficiently move the culture down a specific differentiation pathway.

[0224] Combining a method for isolation and purification of intermediate cell types in the differentiation path is an effective way to deal with the above mentioned problems of heterogeneity and asynchrony.

PRODUCTION OF DEFINITIVE ENDODERM FROM PLURIPOTENT CELLS

[0225] The definitive endoderm cell cultures and compositions comprising definitive endoderm cells that are described herein can be produced from pluripotent cells, such as embryonic stem cells. A preferred method utilizes human embryonic stem cells (hESC) as the starting material for definitive endoderm production. The embryonic stem cells used in this method can be cells that originate from the embryonic inner cell mass or those obtained from embryonic gonadal ridges. Human stem cells can be maintained in culture in a pluripotent state without substantial differentiation using methods that are known in the art. Such methods are described, for example, in US Patent Nos. 5,453,357, 5,670,372, 5,690,926 5,843,780, 6,200,806 and 6,251,671 the disclosures of which are incorporated herein by reference in their entireties.

[0226] In some embodiments of the methods described herein, hESCs are maintained on a feeder layer. Any feeder layer which allows hESCs to be maintained in a pluripotent state can be used in the methods described herein. One commonly used feeder layer for the cultivation of human embryonic stem cells is a layer of mouse fibroblasts. More recently, human fibroblast feeder layers have been developed for use in the cultivation of hESCs (see US Patent Application No. 2002/0072117, the disclosure of which is incorporated herein by reference in its entirety). Alternative embodiments of the methods described herein permit the maintenance of pluripotent hESC without the use of a feeder layer. Such methods have been described in US Patent Application No. 2003/0175956, the disclosure of which is incorporated herein by reference in its entirety.

[0227] The human embryonic stem cells used herein are can be maintained in culture either with or without serum. In some embodiments, serum replacement is used. In other embodiments, serum free culture techniques, such as those described in US Patent

Application No. 2003/0190748, the disclosure of which is incorporated herein by reference in its entirety, are used.

[0228] Stem cells are maintained in culture in a pluripotent state by routine passage until it is desired that they be differentiated into definitive endoderm. In some embodiments, differentiation to definitive endoderm is achieved by providing to the stem cell culture a growth factor of the TGFβ superfamily in an amount sufficient to promote differentiation to definitive endoderm. Growth factors of the TGFβ superfamily which are useful for the production of definitive endoderm are selected from the Nodal/Activin or BMP subgroups. In some embodiments of the differentiation methods described herein, the growth factor is selected from the group consisting of Nodal, Activin A, Activin B and BMP4. Additionally, the growth factor Wnt3a is useful for the production of definitive endoderm cells. In certain embodiments of the present invention, combinations of any of the above-mentioned growth factors can be used.

[0229] With respect to some of the embodiments of differentiation methods described herein, the above-mentioned growth factors are provided to the cells so that the growth factors are present in the cultures at concentrations sufficient to promote differentiation of at least a portion of the stem cells to definitive endoderm. In some embodiments of the present invention, the above-mentioned growth factors are present in the cell culture at a concentration of at least about 10 ng/ml, at least about 25 ng/ml, at least about 50 ng/ml, at least about 75 ng/ml, at least about 100 ng/ml, at least about 200 ng/ml, at least about 300 ng/ml, at least about 400 ng/ml, at least about 500 ng/ml, or at least about 1000 ng/ml.

[0230] In certain embodiments of the present invention, the above-mentioned growth factors are removed from the cell culture subsequent to their addition. For example, the growth factors can be removed within about one day, about two days, about three days, about four days, about five days, about six days, about seven days, about eight days, about nine days or about ten days after their addition. In a preferred embodiment, the growth factors are removed about four days after their addition.

[0231] Cultures of definitive endoderm cells can be grown in medium containing reduced serum or no serum. In certain embodiments of the present invention, serum

concentrations can range from about 0.1% v/v to about 20% v/v. In some embodiments, definitive endoderm cells are grown with serum replacement. In other embodiments, definitive endoderm cells are grown in the presence of B27. In such embodiments, the concentration of B27 supplement can range from about 0.2% v/v to about 20% v/v.

[0232] The progression of the hESC culture to definitive endoderm can be monitored by quantitating expression of marker genes characteristic of definitive endoderm as well as the lack of expression of marker genes characteristic of hESCs and other cell types. One method of quantitating gene expression of such marker genes is through the use of quantitative PCR (Q-PCR). Methods of performing Q-PCR are well known in the art. Other methods which are known in the art can also be used to quantitate marker gene expression. For example, marker gene expression can be detected by using antibodies specific for the marker gene of interest.

[0233] As described further in the Examples below, a reliable marker of definitive endoderm is the SOX17 gene. As such, the definitive endoderm cells produced by the methods described herein express the SOX17 marker gene. Other markers of definitive endoderm are MIXL1, GSC and HNF3b. In some embodiments of the present invention, definitive endoderm cells express the SOX17 marker gene at a level higher than that of the SOX7 marker gene, which is characteristic of visceral endoderm (see Table 1). Additionally, in some embodiments, expression of the SOX17 marker gene is higher than the expression of the OCT4 marker gene, which is characteristic of hESCs. In other embodiments of the present invention, definitive endoderm cells express the SOX17 marker gene at a level higher than the that of the AFP, SPARC or Thrombomodulin (TM) marker genes. In certain embodiments of the present invention, the SOX17-expressing definitive endoderm cells produced by the methods described herein do not express PDX1 (PDX1-negative).

[0234] Another marker of definitive endoderm is the CXCR4 gene. The CXCR4 gene encodes a cell surface chemokine receptor whose ligand is the chemoattractant SDF-1. The principal roles of the CXCR4 receptor-bearing cells in the adult are believed to be the migration of hematopoetic cells to the bone marrow, lymphocyte trafficking and the differentiation of various B cell and macrophage blood cell lineages [Kim, C., and Broxmeyer, H. J. Leukocyte Biol. 65, 6-15 (1999)]. The CXCR4 receptor also functions as

a coreceptor for the entry of HIV-1 into T-cells [Feng, Y., et al. Science, 272, 872-877 (1996)]. In an extensive series of studies carried out by [McGrath, K.E. et al. Dev. Biology 213, 442-456 (1999)], the expression of the chemokine receptor CXCR4 and its unique ligand, SDF-1 [Kim, C., and Broxmyer, H., J. Leukocyte Biol. 65, 6-15 (1999)], were delineated during early development and adult life in the mouse. The CXCR4/SDF1 interaction in development became apparent when it was demonstrated that if either gene was disrupted in transgenic mice [Nagasawa et al. Nature, 382, 635-638 (1996)], Ma, Q., et al Immunity, 10, 463-471 (1999)] it resulted in late embryonic lethality. McGrath et al. demonstrated that CXCR4 is the most abundant chemokine receptor messenger RNA detected during early gastrulating embryos (E7.5) using a combination of RNase protection and in situ hybridization methodologies. In the gastrulating embryo, CXCR4/SDF-1 signaling appears to be mainly involved in inducing migration of primitive-streak germlayer cells and is expressed on definitive endoderm, mesoderm and extraembryonic mesoderm present at this time. In E7.2-7.8 mouse embryos, CXCR4 and alpha-fetoprotein are mutually exclusive indicating a lack of expression in visceral endoderm [McGrath, K.E. et al. Dev. Biology 213, 442-456 (1999)].

[0235] In some embodiments of the present invention, the definitive endoderm cells produced by the methods described herein express the CXCR4 marker gene. In other embodiments, the definitive endoderm cells produced by the methods described herein express the CXCR4 marker gene as well as other markers of definitive endoderm, including but not limited to, SOX17, MIXL1, GSC and HNF3b. In some embodiments of the present invention, definitive endoderm cells express the CXCR4 marker gene at a level higher than that of the SOX7 marker gene. Additionally, in some embodiments, expression of the CXCR4 marker gene is higher than the expression of the OCT4 marker gene. In other embodiments of the present invention, definitive endoderm cells express the CXCR4 marker gene at a level higher than the that of the AFP, SPARC or Thrombomodulin (TM) marker genes. In certain embodiments of the present invention, the CXCR4-expressing definitive endoderm cells produced by the methods described herein do not express PDX1 (PDX1-negative).

[0236] It will be appreciated that expression of CXCR4 in endodermal cells does not preclude the expression of SOX17. Accordingly, in some embodiments of the present invention, definitive endoderm cells are those that express both the SOX17 and CXCR4 marker genes at a level higher than that of the SOX7 marker gene. Additionally, in some embodiments, the expression of both the SOX17 and CXCR4 marker genes is higher than the expression of the OCT4 marker gene. In other embodiments of the present invention, definitive endoderm cells express both the SOX17 and the CXCR4 marker genes at a level higher than the that of the AFP, SPARC or Thrombomodulin (TM) marker genes. In certain embodiments of the present invention, the SOX17/CXCR4-expressing definitive endoderm cells produced by the methods described herein do not express PDX1 (PDX1-negative).

[0237] It will be appreciated that SOX17 and/or CXCR4 marker expression is induced over a range of different levels in definitive endoderm cells depending on the differentiation conditions. As such, in some embodiments of the present invention, the expression of the SOX17 marker and/or the CXCR4 marker in definitive endoderm cells is at least about 2-fold higher to at least about 10,000-fold higher than the expression of the SOX17 marker and/or the CXCR4 marker in undifferentiated cells, such as pluripotent stem cells. In other embodiments of the present invention, the expression of the SOX17 marker and/or the CXCR4 marker in definitive endoderm cells is at least about 4-fold higher, at least about 6-fold higher, at least about 8-fold higher, at least about 10-fold higher, at least about 15-fold higher, at least about 20-fold higher, at least about 40-fold higher, at least about 80-fold higher, at least about 100-fold higher, at least about 150-fold higher, at least about 200-fold higher, at least about 500-fold higher, at least about 750-fold higher, at least about 1000-fold higher, at least about 2500-fold higher, at least about 5000-fold higher, at least about 7500-fold higher or at least about 10,000-fold higher than the expression of the SOX17 marker and/or the CXCR4 marker in undifferentiated cells, such as pluripotent stem cells. In some embodiments, the expression of the SOX17 marker and/or CXCR4 marker in definitive endoderm is infinitely higher than the expression of the SOX17 marker and/or the CXCR4 marker in undifferentiated cells, such as pluripotent stem cells..

[0238] It will also be appreciated that there is a range of differences between the expression level of the SOX17 marker and the expression levels of the AFP, TM and/or

SOX7 markers in definitive endoderm cells. Similarly, there exists a range of differences between the expression level of the CXCR4 marker and the expression levels of the AFP, TM and/or SOX7 markers in definitive endoderm cells. As such, in some embodiments of the present invention, the expression of the SOX17 marker or the CXCR4 marker is at least about 2-fold higher to at least about 10,000-fold higher than the expression of AFP, TM and/or SOX7 markers. In other embodiments of the present invention, the expression of the SOX17 marker or the CXCR4 marker is at least about 4-fold higher, at least about 6-fold higher, at least about 8-fold higher, at least about 10-fold higher, at least about 15-fold higher, at least about 20-fold higher, at least about 40-fold higher, at least about 80-fold higher, at least about 100-fold higher, at least about 150-fold higher, at least about 200-fold higher, at least about 500-fold higher, at least about 750-fold higher, at least about 1000-fold higher, at least about 2500-fold higher, at least about 5000-fold higher, at least about 7500-fold higher or at least about 10,000-fold higher than the expression of AFP, TM and/or SOX7 markers. In some embodiments, AFP, TM and/or SOX7 markers are not detectably expressed in definitive endoderm cells.

COMPOSITIONS COMPRISING DEFINITIVE ENDODERM

[0239] Some aspects of the present invention relate to cell populations or cell cultures that comprise both pluripotent cells, such as stem cells, and definitive endoderm cell. For example, using the methods described herein, compositions comprising mixtures of hESCs and definitive endoderm cells can be produced. In some embodiments, compositions comprising at least about 5 definitive endoderm cells for about every 95 stem cells are produced. In other embodiments, compositions comprising at least about 95 definitive endoderm cells for about every 5 stem cells are produced. Additionally, compositions comprising other ratios of definitive endoderm cells to stem cells are contemplated. For example, compositions comprising at least about 5 definitive endoderm cells for about every 1 stem cell, at least about 10 definitive endoderm cells for about every 1 stem cell, at least about 20 definitive endoderm cells for about every 1 stem cell, at least about 50 definitive endoderm cells for about every 1 stem cell, at least about 100 definitive endoderm cells for about every 1 stem cell, at least about 1000 definitive endoderm cells for about every 1 stem

cell, at least about 10,000 definitive endoderm cells for about every 1 stem cell, at least about 100,000 definitive endoderm cells for about every 1 stem cell and at least about 1,000,000 definitive endoderm cells for about every 1 stem cell are contemplated.

[0240] Some aspects of the present invention relate to mammalian cell compositions comprising from at least about 5% definitive endoderm cells to at least about 95% definitive endoderm cells. Certain specific embodiments relate to mammalian cell compositions comprising from at least about 10% definitive endoderm cells, from at least about 20% definitive endoderm cells, from at least about 30% definitive endoderm cells, from at least about 40% definitive endoderm cells, from at least about 50% definitive endoderm cells, from at least about 60% definitive endoderm cells, from at least about 70% definitive endoderm cells, from at least about 80% definitive endoderm cells, or from at least about 90% definitive endoderm cells.

[0241] Additional embodiments of the present invention relate to compositions comprising mammalian endodermal cells, wherein the expression of either the SOX17 or the CXCR4 marker is greater than the expression of the alpha-fetoprotein (AFP), the Thrombomodulin (TM) and/or the SOX7 marker in at least about 10% of the endodermal cells. In other embodiments, the expression of either the SOX17 or the CXCR4 marker is greater than the expression of the alpha-fetoprotein (AFP), the Thrombomodulin (TM) and/or the SOX7 marker in at least about 20% of the endodermal cells, in at least about 30% of the endodermal cells, in at least about 40% of the endodermal cells, in at least about 50% of the endodermal cells, in at least about 60% of the endodermal cells, in at least about 70% of the endodermal cells, in at least about 80% of the endodermal cells, in at least about 90% of the endodermal cells or in at least about 95% of the endodermal cells.

[0242] Still other embodiments of the present invention relate to compositions comprising mammalian endodermal cells, wherein the expression both the SOX17 and the CXCR4 marker is greater than the expression of the alpha-fetoprotein (AFP), the Thrombomodulin (TM) and/or the SOX7 marker in at least about 10% of the endodermal cells. In other embodiments, the expression of both the SOX17 and the CXCR4 marker is greater than the expression of the alpha-fetoprotein (AFP), the Thrombomodulin (TM) and/or the SOX7 marker in at least about 20% of the endodermal cells, in at least about 30% of the

endodermal cells, in at least about 40% of the endodermal cells, in at least about 50% of the endodermal cells, in at least about 60% of the endodermal cells, in at least about 70% of the endodermal cells, in at least about 80% of the endodermal cells, in at least about 90% of the endodermal cells or in at least about 95% of the endodermal cells.

[0243] Using the methods described herein, compositions comprising definitive endoderm cells substantially free of other cell types can be produced. With respect to cells in cell cultures, the term "substantially free of" means that the specified cell type is present in an amount of less than about 5% of the total number of cells present in the cell culture. In some embodiments of the present invention, the definitive endoderm cell populations or cell cultures produced by the methods described herein are substantially free of cells expressing the OCT4, SOX7, AFP, SPARC, TM, ZIC1 or BRACH marker genes.

[0244] In one embodiment of the present invention, a description of a definitive endoderm cell based on the expression of marker genes is, SOX17 high, MIXL1 high, AFP low, SPARC low, Thrombomodulin low, SOX7 low, CXCR4 high.

ENRICHMENT, ISOLATION AND/OR PURIFICATION OF DEFINITIVE ENDODERM

[0245] With respect to additional aspects of the present invention, definitive endoderm cells can be enriched, isolated and/or purified by using an affinity tag that is specific for such cells. Examples of affinity tags specific for definitive endoderm cells are an antibodies, ligands or other binding agents that are specific to a marker polypeptide that is present on the cell surface of definitive endoderm cells but which is not substantially present on other cell types that would be found in a cell culture produced by the methods described herein. In some embodiments, a SOX17 or CXCR4 antibody is used as an affinity tag for the enrichment, isolation or purification of definitive endoderm cells. In other embodiments, the chemokine SDF-1 or other molecules based on SDF-1 can also be used as affinity tags. Such molecules include, but not limited to, SDF-1 fragments, SDF-1 fusions or SDF-1 mimetics.

[0246] Methods for using antibodies for cell isolation are known in the art and such methods can be implemented for use with the antibodies and cells described herein. In one embodiment, the SOX17 antibody is attached to a magnetic bead then allowed to bind to definitive endoderm cells in a cell culture which has been enzymatically treated to reduce

intercellular and substrate adhesion. The cell/antibody/bead complexes are then exposed to a movable magnetic field which is used to separate bead-bound definitive endoderm cells from unbound cells. Once the definitive endoderm cells are physically separated from other cells in culture, the antibody binding is disrupted and the cells are replated in appropriate tissue culture medium.

[0247] A preferred embodiment for purification of definitive endoderm cells uses an antibody for CXCR4 as an affinity tag. For example, in some embodiments, the CXCR4 antibody is incubated with definitive endoderm-containing cell culture that has been treated to reduce intercellular and substrate adhesion. The cells are then washed, centrifuged and resuspended. The cell suspension is then incubated with a secondary antibody, such as an FITC-conjugated antibody that capable of binding to the primary antibody. The cells are then washed, centrifuged and resuspended in buffer. The cell suspension is then analyzed and sorted using a fluorescence activated cell sorter (FACS). CXCR4-positive cells are collected separately from CXCR4-negative cells, thereby resulting in the isolation of such cell types. If desired, the isolated cell compositions can be further purified by using an alternate affinity-based method or by additional rounds of sorting using the same or different markers that are specific for definitive endoderm.

[0248] In other embodiments of the present invention, definitive endoderm is enriched, isolated and/or purified using a ligand or other molecule that binds to CXCR4. In some embodiments, the molecule is SDF-1 or a fragment, fusion or mimetic thereof.

[0249] In preferred embodiments, definitive endoderm cells are enriched, isolated or purified from other non-definitive endoderm cells after the stem cell cultures are induced to differentiate towards the definitive endoderm lineage. It will be appreciated that the above-described enrichment, isolation and purification procedures can be used with such cultures at any stage of differentiation.

[0250] In addition to the procedures just described, definitive endoderm cells may also be isolated by other techniques for cell isolation. Additionally, definitive endoderm cells may also be enriched or isolated by methods of serial subculture in growth conditions which promote the selective survival or selective expansion of said definitive endoderm cells.

[0251] Using the methods described herein, enriched, isolated and/or purified populations of definitive endoderm cells and or tissues can be produced *in vitro* from pluripotent cell cultures or cell populations, such as stem cell cultures or populations, which have undergone at least some differentiation. In some embodiments, the cells undergo random differentiation. In a preferred embodiment, however, the cells are directed to differentiate primarily into definitive endoderm. Some preferred enrichment, isolation and/or purification methods relate to the *in vitro* production of definitive endoderm from human embryonic stem cells. Using the methods described herein, cell populations or cell cultures can be enriched in definitive endoderm content by at least about 2- to about 1000-fold as compared to untreated cell populations or cell cultures. In some embodiments, definitive endoderm cells can be enriched by at least about 5- to about 500-fold as compared to untreated cell populations or cell cultures. In other embodiments, definitive endoderm cells can be enriched from at least about 10- to about 200-fold as compared to untreated cell populations or cell cultures. In still other embodiments, definitive endoderm cells can be enriched from at least about 20- to about 100-fold as compared to untreated cell populations or cell cultures. In yet other embodiments, definitive endoderm cells can be enriched from at least about 40- to about 80-fold as compared to untreated cell populations or cell cultures. In certain embodiments, definitive endoderm cells can be enriched from at least about 2- to about 20-fold as compared to untreated cell populations or cell cultures.

[0252] Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only, and are not intended to be limiting.

EXAMPLES

[0253] Many of the examples below describe the use of human embryonic stem cells. Methods of producing human embryonic stem cells are well known in the art and have been described numerous scientific publications.

EXAMPLE 1

Human ES cells

[0254] For our studies of endoderm development we employed human embryonic stem cells, which are pluripotent and can divide seemingly indefinitely in culture while maintaining a normal karyotype. ES cells were derived from the 5-day-old embryo inner cell mass using either immunological or mechanical methods for isolation. The hESC line designated hESCyT-25 was derived more than 18 months ago and has been serially passaged over 50 times. We employed the hESCyT-25 human embryonic stem cell line as our starting material for the production of definitive endoderm.

[0255] It will be appreciated by those of skill in the art that stem cells or other pluripotent cells can also be used as starting material for the differentiation procedures described herein. For example, cells obtained from embryonic gonadal ridges, which can be isolated by methods known in the art, can be used as pluripotent cellular starting material.

EXAMPLE 2

hESCyt-25 Characterization

[0256] The human embryonic stem cell line, hESCyt-25 has maintained a normal morphology, karyotype, growth and self-renewal properties over 18 months in culture as shown in Figures 2A-B and 3. hESCyt-25 displays strong immunoreactivity for the OCT4, SSEA-4 and TRA-1-60 antigens characteristic of undifferentiated hESCs (Figures 4A-C) and displays alkaline phosphatase activity as well as a morphology identical to other established hESC lines (Figures 5A-D). The hESCyT-25 cell line also readily forms embryoid bodies (EBs) when cultured in suspension as shown in Figure 6. As a demonstration of its pluripotent nature, hESCyT-25 differentiates into various cell types that represent the three principal germ layers. Ectoderm production was demonstrated by Q-PCR for ZIC1 as well as immunocytochemistry (ICC) for nestin and more mature neuronal markers. Immunocytochemical staining for β-III tubulin was observed in clusters of elongated cells, characteristic of early neurons (Figure 7). Previously, we treated EBs in suspension with retinoic acid, to induce differentiation of pluripotent stem cells to visceral endoderm (VE), an extra-embryonic lineage. Treated cells expressed high levels of α-fetoprotein (AFP) and SOX7, two markers of VE, by 54 hours of treatment (Figure 8A). Cells differentiated in monolayer expressed AFP in sporadic patches as demonstrated by immunocytochemical

staining (Figure 8B). As will be described below, the hESCyT-25 cell line was also capable of forming definitive endoderm, as validated by real-time quantitative polymerase chain reaction (Q-PCR) and immunocytochemistry for SOX17, in the absence of AFP expression. To demonstrate differentiation to mesoderm, differentiating EBs were analyzed for Brachyury gene expression at several time points. Brachyury expression increased progressively over the course of the experiment (Figure 9). In view of the foregoing, the hESCyT-25 line is pluripotent as shown by the ability to form cells representing the three germ layers.

EXAMPLE 3

Production of SOX17 Antibody

[0257] A primary obstacle to the identification of definitive endoderm in hESC cultures is the lack of appropriate tools. We therefore undertook the production of an antibody raised against human SOX17 protein.

[0258] SOX17 is expressed throughout the definitive endoderm as it forms during gastrulation and its expression is maintained in the gut tube (although levels of expression vary along the A-P axis) until around the onset of organogenesis. SOX17 is also expressed in a subset of extra-embryonic endoderm cells. No expression of this protein has been observed in mesoderm or ectoderm. As such, SOX17 is an appropriate marker for the definitive endoderm lineage when used in conjunction with markers to exclude extra-embryonic lineages.

[0259] As described in detail herein, the SOX17 antibody was utilized to specifically examine effects of various treatments and differentiation procedures aimed at the production of SOX17 positive definitive endoderm cells. Other antibodies reactive to AFP, SPARC and Thrombomodulin were also employed to rule out the production of visceral and parietal endoderm (extra-embryonic endoderm).

[0260] A portion of the human SOX17 cDNA (SEQ ID NO: 1) corresponding to amino acids 172-414 in the carboxyterminal end of the SOX17 protein (SEQ ID NO: 2) was used for production of SOX17 antibody by genetic immunization (Figure 10). The SOX17 antibody was produced by genetic immunization in rats at the antibody production company, GENOVAC (Freiberg, Germany), according to procedures developed there. Procedures for

genetic immunization can be found in US Patent Nos. 5,830,876, 5,817,637, 6,165,993 and 6,261,281 as well as International Patent Application Publication Nos. WO00/29442 and WO99/13915, the disclosures of which are incorporated herein by reference in their entireties.

[0261] Other suitable methods for genetic immunization are also described in the non-patent-related literature. For example, Barry et al. describe the production of monoclonal antibodies by genetic immunization in *Biotechniques* **16**: 616-620, 1994, the disclosure of which is incorporated herein by reference in its entirety. Specific examples of genetic immunization methods to produce antibodies against specific proteins can be found, for example, in Costaglia *et al.*, (1998) Genetic immunization against the human thyrotropin receptor causes thyroiditis and allows production of monoclonal antibodies recognizing the native receptor. *J. Immunol.* 160: 1458-1465; Kilpatrick *et al* (1998) Gene gun delivered DNA-based immunizations mediate rapid production of murine monoclonal antibodies to the Flt-3 receptor. *Hybridoma* 17: 569-576; Schmolke *et al.*, (1998) Identification of hepatitis G virus particles in human serum by E2-specific monoclonal antibodies generated by DNA immunization. *J. Virol.* 72: 4541-4545; Krasemann et al., (1999) Generation of monoclonal antibodies against proteins with an unconventional nucleic acid-based immunization strategy. *J. Biotechnol.* 73: 119-129; and Ulivieri *et al.*, (1996) Generation of a monoclonal antibody to a defined portion of the *Heliobacter pylori* vacuolating cytotoxin by DNA immunization. *J. Biotechnol.* 51: 191-194, the disclosures of which are incorporated herein by reference in their entireties.

[0262] The antibody produced by genetic immunization was determined to be specific for SOX17 by both Western blot and ICC on hSOX17 cDNA-transfected cell lines as shown in Figures 11 and 12. SOX7 and SOX18 are the closest Sox family relatives to SOX17 as depicted in the relational dendrogram shown in Figure 13. We employed the human SOX7 polypeptide as a negative control to demonstrate that the SOX17 antibody is specific for SOX17 and does not react with its closest family member.

[0263] The following methods were utilized for the production of the SOX17 and SOX7 expression vectors, their transfection into human fibroblasts and analysis by Western blot. Expression vectors employed for the production of SOX17, SOX7, and EGFP were

pCMV6 (OriGene Technologies, Inc., Rockville, MD), pCMV-SPORT6 (Invitrogen, Carlsbad, CA) and pEGFP-N1 (Clonetech, Palo Alto, CA), respectively. For protein production, telomerase immortalized MDX human fibroblasts were transiently transfected with supercoiled DNA in the presence of Lipofectamine 2000 (Invitrogen, Carlsbad, CA). Total cellular lysates were collected 36 hours post-transfection in 50 mM TRIS-HCl (pH 8), 150 mM NaCl, 0.1% SDS, 0.5% deoxycholate, containing a cocktail of protease inhibitors (Roche Diagnostics Corporation, Indianapolis, IN). Western blot analysis of 100 µg of cellular proteins, separated by SDS-PAGE on NuPAGE (4-12 % gradient polyacrylamide, Invitrogen, Carlsbad, CA), and transferred by electro-blotting onto PDVF membranes (Hercules, CA), were probed with a 1/1000 dilution of the rat SOX17 anti-serum in 10 mM TRIS-HCl (pH 8), 150 mM NaCl, 10% BSA, 0.05 % Tween-20 (Sigma, St. Louis, MO), followed by Alkaline Phosphatase conjugated anti-rat IgG (Jackson ImmunoResearch Laboratories, West Grove, PA), and revealed through Vector Black Alkaline Phosphatase staining (Vector Laboratories, Burlingame, CA). The proteins size standard used was wide range color markers (Sigma, St. Louis, MO).

[0264] In Figure 11, protein extracts made from human fibroblast cells that were transiently transfected with SOX17, SOX7 or EGFP cDNA's were probed on Western blots with the SOX17 antibody. Only the protein extract from hSOX17 transfected cells produced a band of ~51Kda which closely matched the predicted 46 Kda molecular weight of the human SOX17 protein. There was no reactivity of the SOX17 antibody to extracts made from either human SOX7 or EGFP transfected cells. Furthermore, the SOX17 antibody clearly labeled the nuclei of human fibroblast cells transfected with the hSOX17 expression construct (Figures 12A-C) but did not label cells transfected with EGFP alone (Figures 12D-F). As such, the SOX17 antibody exhibits specificity by ICC (Figures 12A-F).

EXAMPLE 4

Validation of SOX17 Antibody as a Marker of Definitive Endoderm

[0265] As evidence that the SOX17 antibody is specific for human SOX17 protein and furthermore marks definitive endoderm, partially differentiated hESCs were co-labeled with SOX17 and AFP antibodies. It has been previously demonstrated that SOX17,

SOX7, which is a closely related member of the SOX gene family subgroup F (Figure 13), and AFP are each expressed in visceral endoderm. However, AFP and SOX7 are not expressed in definitive endoderm cells at levels detectable by ICC, and thus, they be employed as negative markers for bonifide definitive endoderm cells. As shown in Figures 14A-B, SOX17 antibody labels populations of cells that exist as discrete groupings of cells or are intermingled with AFP positive cells. In particular, Figure 14A shows that small numbers of SOX17 cells were co-labeled with AFP, which are consistent with the results of *in vivo* studies (Kanai-Azuma et al.) where a subset of visceral endoderm cells was shown to express SOX17 mRNA. However, regions were also found where there were little or no AFP^+ cells in the field of $SOX17^+$ cells (Figure 14B). Similarly, since parietal endoderm has also been reported to express SOX17, antibody co-labeling with SOX17 together with the parietal markers Sparc and/or Thrombomodulin (TM) can be used to identify the $SOX17^+$ cells which are parietal endoderm. As shown in Figures 15A-C, Thrombomodulin and SOX17 co-labelled parietal endoderm cells were produced by random differentiation of hES cells.

[0266] In view of the above cell labelling experiments, the identity of a definitive endoderm cell can be established by the marker profile $SOX17^{hi}/AFP^{lo}/[TM^{lo}$ or $SPARC^{lo}]$. In other words, the expression of the SOX17 marker is greater than the expression of the AFP marker, which is characteristic of visceral endoderm, and the TM or SPARC markers, which are characteristic of parietal endoderm. Accordingly, those cells positive for SOX17 but negative for AFP and negative for TM or SPARC are definitive endoderm.

[0267] As a further evidence of the specificity of the $SOX17^{hi}/AFP^{lo}/TM^{lo}/SPARC^{lo}$ marker profile as predictive of definitive endoderm, SOX17 and AFP gene expression was quantitatively compared to the relative number of antibody labeled cells. As shown in Figure 16A, hESCs treated with retinoic acid (visceral endoderm inducer), or Activin A (definitive endoderm inducer), resulted in a 10-fold difference in the level of SOX17 mRNA expression. This result mirrored the 10-fold difference in SOX17 antibody-labeled cell number (Figure 16B). Furthermore, as shown in Figure 17A, Activin A treatment of hESCs suppressed AFP gene expression by 6.8-fold in comparison to no treatment. This was visually reflected by a dramatic decrease in the number of AFP labeled cells in these cultures as shown in Figure 17B. To quantify this further it was demonstrated

that this approximately 7-fold decrease in AFP gene expression was the result of a similar 7-fold decrease in AFP antibody-labeled cell number as measured by flow cytometry (Figure 18). This result is extremely significant in that it indicates that quantitative changes in gene expression as seen by Q-PCR mirror changes in cell type specification as observed by antibody staining.

[0268] Incubation of hESCs in the presence of Nodal family members (Nodal, Activin A and Activin B - NAA) resulted in a significant increase in SOX17 antibody-labeled cells over time. By 5 days of continuous activin treatment greater than 50% of the cells were labeled with SOX17 (Figures 19A-F). There were few or no cells labeled with AFP after 5 days of activin treatment.

[0269] In summary, the antibody produced against the carboxy-terminal 242 amino acids of the human SOX17 protein identifies human SOX17 protein on Western blots but does not recognize SOX7, it's closest Sox family relative. The SOX17 antibody recognizes a subset of cells in differentiating hESC cultures that are primarily $SOX17^{+}/AFP^{lo/-}$ (greater than 95% of labeled cells) as well as a small percentage (< 5%) of cells that co-label for SOX17 and AFP (visceral endoderm). Treatment of hESC cultures with activins results in a marked elevation of SOX17 gene expression as well as SOX17 labeled cells and dramatically suppresses the expression of AFP mRNA and the number of cells labeled with AFP antibody.

EXAMPLE 5

Q-PCR Gene Expression Assay

[0270] Real-time quantitative RT-PCR (Q-PCR) was the primary assay used here for screening the effects of various treatments on hESC differentiation. Real-time measurements of gene expression were analyzed for multiple marker genes at multiple time points by Q-PCR. Marker genes characteristic of the desired as well as undesired cell types were evaluated to gain a better understanding of the overall dynamics of the cellular populations. The strength of Q-PCR analysis includes its extreme sensitivity and relative ease of developing the necessary markers, as the genome sequence is readily available. Furthermore, the extremely high sensitivity of Q-PCR permits detection of gene expression

from a relatively small number of cells within a much larger population. In addition, the ability to detect very low levels of gene expression provides indications for "differentiation bias" within the population. The bias towards a particular differentiation pathway, prior to the overt differentiation of those cellular phenotypes, is unrecognizable using immunocytochemical techniques. For this reason, Q-PCR provides a method of analysis that is at least complementary and potentially much superior to immunocytochemical techniques for screening the success of differentiation treatments. Additionally, Q-PCR provides a mechanism by which to evaluate the success of a differentiation protocol in a quantitative format at semi-high throughput scales of analysis.

[0271] The approach taken here was to perform relative quantitation using SYBR Green chemistry on the Rotor Gene 3000 instrument (Corbett Research) and a two-step RT-PCR format. Such an approach allows for the banking of cDNA samples for analysis of additional marker genes in the future, thus avoiding variability in the reverse transcription efficiency between samples.

[0272] Primers were designed to lie over exon-exon boundaries or span introns of at least 800 bp when possible, as this has been empirically determined to eliminate amplification from contaminating genomic DNA. When marker genes were employed that do not contain introns or they possess pseudogenes, DNase I treatment of RNA samples was performed.

[0273] We routinely used Q-PCR to measure the gene expression of multiple markers of target and non-target cell types in order to provide a broad profile description of gene expression in cell samples. The markers relevant for the early phases of hESC differentiation (specifically ectoderm, mesoderm, definitive endoderm and extra-embryonic endoderm) and for which validated primer sets are available are provided below in Table 1. The human specificity of these primer sets has also been demonstrated. This is an important fact since the hESCs were often grown on mouse feeder layers. Most typically, triplicate samples were taken for each condition and independently analyzed in duplicate to assess the biological variability associated with each quantitative determination.

[0274] To generate PCR template, total RNA was isolated using RNeasy (Qiagen) and quantitated using RiboGreen (Molecular Probes). Reverse transcription from 350-500 ng

of total RNA was carried out using the iScript reverse transcriptase kit (BioRad), which contains a mix of oligo-dT and random primers. Each 20 μL reaction was subsequently diluted up to 100 μL total volume and 3 μL was used in each 10 μL Q-PCR reaction containing 400 nM forward and reverse primers and 5 μL 2X SYBR Green master mix (Qiagen). Two step cycling parameters were used employing a 5 second denature at 85-94°C (specifically selected according to the melting temp of the amplicon for each primer set) followed by a 45 second anneal/extend at 60°C. Fluorescence data was collected during the last 15 seconds of each extension phase. A three point, 10-fold dilution series was used to generate the standard curve for each run and cycle thresholds (Ct's) were converted to quantitative values based on this standard curve (Figure 19). The quantitated values for each sample were normalized to housekeeping gene performance and then average and standard deviations were calculated for triplicate samples. At the conclusion of PCR cycling, a melt curve analysis was performed to ascertain the specificity of the reaction. A single specific product was indicated by a single peak at the T_m appropriate for that PCR amplicon (Figure 20). In addition, reactions performed without reverse transcriptase served as the negative control and do not amplify.

[0275] A first step in establishing the Q-PCR methodology was validation of appropriate housekeeping genes (HGs) in the experimental system. Since the HG was used to normalize across samples for the RNA input, RNA integrity and RT efficiency, it was of value that the HG exhibited a constant level of expression over time in all sample types in order for the normalization to be meaningful. We measured the expression levels of *Cyclophilin G, hypoxanthine phosphoribosyltransferase 1 (HPRT), beta-2-microglobulin, hydroxymethylbiane synthase (HMBS), TATA-binding protein (TBP), and glucoronidase beta (GUS)* in differentiating hESCs. Our results indicated that *beta-2-microglobulin* expression levels increased over the course of differentiation and therefore we excluded the use of this gene for normalization. The other genes exhibited consistent expression levels over time as well as across treatments. We routinely used both Cyclophilin G and GUS to calculate a normalization factor for all samples. The use of multiple HGs simultaneously reduces the variability inherent to the normalization process and increases the reliability of the relative gene expression values (Vandesompele et al., 2002).

[0276] After obtaining genes for use in normalization, Q-PCR was then utilized to determine the relative gene expression levels of many marker genes across samples receiving different experimental treatments. The marker genes employed have been chosen because they exhibit enrichment in specific populations representative of the early germ layers and in particular have focused on sets of genes that are differentially expressed in definitive endoderm and extra-embryonic endoderm. These genes as well as their relative enrichment profiles are highlighted in Table 1.

TABLE 1

Germ Layer	Gene	Expression Domains
Endoderm	SOX17	definitive, visceral and parietal endoderm
	MIXL1	endoderm and mesoderm
	GATA4	definitive and primitive endoderm
	HNF3b	definitive endoderm and primitive endoderm, mesoderm, neural plate
Extra-embryonic	SOX7	visceral endoderm
	AFP	visceral endoderm, liver
	SPARC	parietal endoderm
	TM	parietal endoderm/trophectoderm
Ectoderm	ZIC1	neural tube, neural progenitors
Mesoderm	BRACH	nascent mesoderm

[0277] Since many genes are expressed in more than one germ layer it is useful to quantitatively compare expression levels of many genes within the same experiment. SOX17 is expressed in definitive endoderm and to a smaller extent in visceral and parietal endoderm, SOX7 and AFP are expressed in visceral endoderm at this early developmental time point, SPARC, and TM are expressed in parietal endoderm and Brachyury is expressed in early mesoderm.

[0278] Definitive endoderm cells were predicted to express high levels of SOX17 mRNA and low levels of AFP and SOX7 (visceral endoderm), SPARC (parietal endoderm) and Brachyury (mesoderm). In addition, ZIC1 was used here to further rule out induction of early ectoderm. Finally, GATA4 and HNF3b were expressed in both definitive and extra-embryonic endoderm, and thus, correlate with SOX17 expression in definitive endoderm (Table 1). A representative experiment is shown in Figures 21-24 which demonstrates how the marker genes described in Table 1 correlate with each other among the various samples, thus highlighting specific patterns of differentiation to definitive endoderm and extra-embryonic endoderm as well as to mesodermal and neural cell types.

[0279] In view of the above data it is clear that increasing doses of activin results in increasing SOX17 gene expression. Further this SOX17 expression appears to predominantly represent definitive endoderm as opposed to extra-embryonic endoderm. This

conclusion stems from the observation that SOX17 gene expression is inversely correlated with AFP, SOX7, and SPARC gene expression.

EXAMPLE 6

Directed Differentiation of Human ES Cells to Definitive Endoderm

[0280] Human ES cell cultures will randomly differentiate if they are cultured under conditions that do not actively maintain their undifferentiated state. This heterogeneous differentiation results in production of extra-embryonic endoderm cells comprised of both parietal and visceral endoderm (AFP, SPARC and SOX7 expression) as well as early ectodermal and mesodermal derivatives as marked by ZIC1 and Nestin (ectoderm) and Brachyury (mesoderm) expression. Definitive endoderm cell appearance has not traditionally been examined or specified for lack of specific antibody markers in ES cell cultures. As such and by default, early definitive endoderm production in ES cell cultures has not been well studied. Since no good antibody reagents for definitive endoderm cells have been available, most of the characterization has focused on ectoderm and extra-embryonic endoderm. Overall, there are significantly greater numbers of extra-embryonic and neurectodermal cell types in comparison to $SOX17^{hi}$ definitive endoderm cells in randomly differentiated ES cell cultures.

[0281] As undifferentiated hESC colonies expand on a bed of fibroblast feeders the edges of the colony take on alternative morphologies that are distinct from those cells residing within the interior of the colony. Many of these outer edge cells can be distinguished by their less uniform, larger cell body morphology and by the expression of higher levels of OCT4 (Figures 25A-C). It has been described that as ES cells begin to differentiate they alter the levels of OCT4 expression up or down relative to undifferentiated ES cells. Alteration of OCT4 levels above or below the undifferentiated threshold may signify the initial stages of differentiation away from the pluripotent state.

[0282] When undifferentiated colonies were examined by SOX17 immunocytochemistry, occasionally small 10-15-cell clusters of SOX17-positive cells were detected at random locations on the periphery and at the junctions between undifferentiated ESC colonies (Figures 26A&D). As noted above, these scattered pockets of outer colony

edges appeared to be some of the first cells to differentiate away from the classical ESC morphology as the colony expanded in size and became more crowded. Younger, smaller fully undifferentiated colonies (< 1mm; 4-5 days old) showed no SOX17 positive cells within or at the edges of the colonies while older, larger colonies (1-2 mm diameter, > 5days old) had sporadic isolated patches of SOX17 positive, AFP negative cells at the periphery of some colonies or in regions interior to the edge (Figure 26) that were differentiated away from classical hESC morphology shown previously (Figure 2). Given that this was the first development of an effective SOX17 antibody, definitive endoderm cells generated in such early "undifferentiated" ESC cultures have never been previously demonstrated.

[0283] Based on negative correlations of SOX17 and SPARC gene expression levels by Q-PCR, the vast majority of these SOX17 positive, AFP negative cells will be negative for parietal markers by antibody co-labeling. This was specifically demonstrated for TM expressing parietal endoderm cells as shown in Figures 27A-B. Exposure to Nodal factors Activin A and B resulted in a dramatic decrease in the intensity to TM expression and the number of TM positive cells. By triple labeling using SOX17, AFP and TM antibodies on an activin treated culture, many clusters of SOX17 positive cells which were also negative for AFP and TM were observed (Figure 37). These are the first cellular demonstrations of SOX17 positive definitive endoderm cells in differentiating ESC cultures (Figures 37 & 28).

[0284] With the SOX17 antibody and Q-PCR tools described above we have explored a number of procedures capable of efficiently programming ESCs to become $SOX17^{hi}/AFP^{lo}$ / $SPARC/TM^{lo}$ definitive endoderm cells (Figure 28). We applied a variety of differentiation protocols aimed at increasing the number and proliferative capacity of these cells as measured at the population level by Q-PCR for SOX17 gene expression and at the level of individual cells by antibody labeling of SOX17 protein.

[0285] As described earlier, the Nodal/activin subgroup of the TGFβ superfamily of factors (Figure 29) has been investigated with regard to mesoderm and endoderm induction in vivo. It is now generally believed that the Nodal-related TGFβ family member signaling molecules and their receptors are responsible for the induction of mesendoderm i.e. mesoderm and definitive endoderm during gastrulation. However, little data has been

generated analyzing Nodal/activin/BMP effects in vitro with regard to identification of definitive endoderm cells as created from embryonic stem cells.

[0286] As shown in Figure 30, addition of Activin A at 100 ng/ml resulted in a 19-fold induction of SOX17 gene expression vs. undifferentiated hESCs by day 4 of differentiation. Adding Activin B, a second member of the activin family, together with Activin A, resulted in a 37-fold induction over undifferentiated hESCs by day 4 of combined activin treatment. Finally, adding a third member of the TGFβ family from the Nodal/Activin and BMP subgroups, BMP4, together with Activin A and Activin B, increased the fold induction to 57 times that of undifferentiated hESCs (Figure 30). When SOX17 induction with activins and BMP was compared to no factor medium controls 5-, 10-, and 15-fold inductions resulted at the 4-day time point. By five days of triple treatment with Activins A, B and BMP, SOX17 was induced more than 70 times higher than hESCs. These data indicate that higher doses and longer treatment times of the Nodal/activin TGFβ family members results in increased expression of SOX17.

[0287] *In vivo*, mesendoderm gives rise to endoderm and mesoderm. The relative proportions of endoderm and mesoderm are dependent on the level of Nodal signaling as controlled by dose and length of exposure. Higher levels of nodal signaling result in more endoderm at the expense of mesoderm and vice-versa for lower Nodal signaling. However, no data exists showing that the combination of Nodal and related molecules Activin A, B and BMP facilitates the expression of SOX17 and definitive endoderm formation *in vivo* or *in vitro*. Furthermore, addition of BMP results in an improved SOX17 induction possibly through the further induction of Cripto, the Nodal co-receptor.

[0288] Again, the combination of Activins A and B together with BMP4 have not previously been shown to result in additive increases in SOX17 induction and hence definitive endoderm formation. BMP4 addition for prolonged periods (>4 days), in combination with Activin A and B may induce SOX17 in parietal and visceral endoderm as well as definitive endoderm (through Cripto receptor induction). In some embodiments of the present invention, it is therefore valuable to remove BMP4 from the treatment within 4 days of addition.

[0289] To determine the effect of TGFβ factor treatment at the individual cell level, a time course of TGFβ factor addition was examined using SOX17 antibody labeling. As previously shown in Figures 19A-F, there was a dramatic increase in the relative number of SOX17 labeled cells over time. The relative quantification (Figure 31) shows more than a 20-fold increase in SOX17-labeled cells. This result indicates that both the numbers of cells as well SOX17 gene expression level are increasing with time of TGFβ factor exposure. As shown in Figure 32, after four days of exposure to Nodal, Activin A, Activin B and BMP4, the level of SOX17 induction reached 168-fold over undifferentiated hESCs. Figure 33 shows that the relative number of SOX17-positive cells was also dose responsive. Activin A doses of 100 ng/mL or more were capable of potently inducing SOX17 gene expression and cell number.

[0290] In addition to the TGFβ family members, the Wnt family of molecules may play a role in specification and/or maintenance of definitive endoderm. The use of Wnt molecules was also beneficial for the differentiation of hESCs to definitive endoderm as indicted by the increased SOX17 gene expression in samples that were treated with activins plus Wnt3a over that of activins alone (Figure 34).

[0291] All of the experiments described above were performed using tissue culture medium containing 10% serum with added factors. Interestingly, the concentration of serum had an effect on the level of SOX17 expression in the presence of added activins as shown in Figures 35A-C. When serum levels were reduced from 10% to 2%, SOX17 expression tripled in the presence of Activins A and B.

[0292] Finally, we demonstrated that activin induced $SOX17^+$ cells divide in culture as depicted in Figures 36A-D. The arrows show cells labeled with SOX17/PCNA/DAPI that are in mitosis as evidenced by the PCNA/DAPI-labeled mitotic plate pattern and the phase contrast mitotic profile.

EXAMPLE 7

Chemokine receptor 4 (CXCR4) expression correlates with markers for definitive endoderm and not markers for mesoderm, ectoderm or visceral endoderm

[0293] As described above, ESCs can be induced to differentiate to the definitive endoderm germ layer by the application of cytokines of the TGFβ family and more specifically of the activin/nodal subfamily. Additionally, we have shown that the proportion of fetal bovine serum (FBS) in the differentiation culture medium effects the efficiency of definitive endoderm differentiation from ESCs. This effect is such that at a given concentration of activin A in the medium, higher levels of FBS will inhibit maximal differentiation to definitive endoderm. In the absence of exogenous activin A, differentiation of ESCs to the definitive endoderm lineage is very inefficient and the FBS concentration has much milder effects on the differentiation process of ESCs.

[0294] hESCs were differentiated by growing in RPMI medium (Invitrogen, Carlsbad, CA; cat# 61870-036) supplemented with 0.5%, 2.0% or 10% FBS and either with or without 100 ng/mL activin A for 6 days. In addition, a gradient ranging from 0.5% to 2.0% over the first three days of differentiation was also used in conjunction with 100 ng/mL of activin A. After the 6 days, replicate samples were collected from each culture condition and analyzed for relative gene expression by real-time quantitative PCR. The remaining cells were fixed for immunofluorescent detection of SOX17 protein.

[0295] The expression levels of CXCR4 varied dramatically across the 7 culture conditions used (Figure 38). In general, CXCR4 expression was high in activin A treated cultures (A100) and low in those which did not receive exogenous activin A (NF). In addition, among the A100 treated cultures, CXCR4 expression was highest when FBS concentration was lowest. There was a remarkable decrease in CXCR4 level in the 10% FBS condition such that the relative expression was more in line with the conditions that did not receive activin A (NF).

[0296] As described above, expression of the SOX17, GSC, MIXL1, and HNF3β is consistent with the characterization of a cell as definitive endoderm. The relative expression of these four genes across the 7 differentiation conditions mirrors that of CXCR4 (Figures 39A-D). This demonstrates that CXCR4 is also a marker of definitive endoderm.

[0297] Ectoderm and mesoderm lineages can be distinguished from definitive endoderm by their expression of various markers. Early mesoderm expresses the genes Bracyury and MOX1 while nascent neuro-ectoderm expresses SOX1 and ZIC1. Figures 40A-D demonstrate that the cultures which did not receive exogenous activin A were preferentially enriched for mesoderm and ectoderm gene expression and that among the activin A treated cultures, the 10% FBS condition also had increased levels of mesoderm and ectoderm marker expression. These patterns of expression were inverse to that of CXCR4 and indicated that CXCR4 was not highly expressed in mesoderm or ectoderm derived from ESCs at this developmental time period.

[0298] Early during mammalian development, differentiation to extra-embryonic lineages also occurs. Of particular relevance here is the differentiation of visceral endoderm that shares the expression of many genes in common with definitive endoderm, including SOX17. To distinguish definitive endoderm from extra-embryonic visceral endoderm one should examine a marker that is distinct between these two. SOX7 represents a marker that is expressed in the visceral endoderm but not in the definitive endoderm lineage. Thus, culture conditions that exhibit robust SOX17 gene expression in the absence of SOX7 expression are likely to contain definitive and not visceral endoderm. It is shown in Figure 40E that SOX7 was highly expressed in cultures that did not receive activin A, SOX7 also exhibited increased expression even in the presence of activin A when FBS was included at 10%. This pattern is the inverse of the CXCR4 expression pattern and suggests that CXCR4 is not highly expressed in visceral endoderm.

[0299] The relative number of SOX17 immunoreactive ($SOX17^+$) cells present in each of the differentiation conditions mentioned above was also determined. When hESCs were differentiated in the presence of high dose activin A and low FBS concentration (0.5% - 2.0%) $SOX17^+$ cells were ubiquitously distributed throughout the culture. When high dose activin A was used but FBS was included at 10% (v/v), the $SOX17^+$ cells appeared at much lower frequency and always appeared in isolated clusters rather than evenly distributed throughout the culture (Figures 41A&B). A further decrease in $SOX17^+$ cells was seen when no exogenous activin A was used. Under these conditions the $SOX17^+$ cells also appeared in clusters and these clusters were smaller and much more rare than those found in the high

activin A, low FBS treatment (Figure 41C). These results demonstrate that the CXCR4 expression patterns not only correspond to definitive endoderm gene expression but also to the number of committed definitive endoderm cells in each condition.

EXAMPLE 8

Differentiation conditions that enrich for definitive endoderm increase the proportion of CXCR4 positive cells

[0300] The dose of activin A also effects the efficiency at which definitive endoderm can be derived from ESCs. This example demonstrates that increasing the dose of activin A increases the proportion of $CXCR4^+$ cells in the culture.

[0301] hESCs were differentiated in RPMI media supplemented with 0.5%-2% FBS (increased from 0.5% to 1.0% to 2.0% over the first 3 days of differentiation) and either 0, 10, or 100 ng/mL of activin A. After 7 days of differentiation the cells were dissociated in PBS without Ca^{2+}/Mg^{2+} containing 2% FBS and 2 mM (EDTA) for 5 minutes at room temperature. The cells were filtered through 35 um nylon filters, counted and pelleted. Pellets were resuspended in a small volume of 50% human serum/50% normal donkey serum and incubated for 2 minutes on ice to block non-specific antibody binding sites. To this, 1 uL of mouse anti-CXCR4 antibody (Abcam, cat# ab10403-100) was added per 50 uL (containing approximately 10^5 cells) and labeling proceeded for 45 minutes on ice. Cells were washed by adding 5 mL of PBS containing 2% human serum (buffer) and pelleted. A second wash with 5 mL of buffer was completed then cells were resuspended in 50 uL buffer per 10^5 cells. Secondary antibody (FITC conjugated donkey anti-mouse; Jackson ImmunoResearch, cat# 715-096-151) was added at 5 ug/mL final concentration and allowed to label for 30 minutes followed by two washes in buffer as above. Cells were resuspended at $5x10^6$ cells/mL in buffer and analyzed and sorted using a FACS Vantage (Beckton Dickenson) by the staff at the flow cytometry core facility (The Scripps Research Institute). Cells were collected directly into RLT lysis buffer (Qiagen) for subsequent isolation of total RNA for gene expression analysis by real-time quantitative PCR.

[0302] The number of $CXCR4^+$ cells as determined by flow cytometry were observed to increase dramatically as the dose of activin A was increased in the differentiation

culture media (Figures 42A-C). The $CXCR4^+$ cells were those falling within the R4 gate and this gate was set using a secondary antibody-only control for which 0.2% of events were located in the R4 gate. The dramatically increased numbers of $CXCR4^+$ cells correlates with a robust increase in definitive endoderm gene expression as activin A dose is increased (Figures 44A-D).

EXAMPLE 9

Isolation of CXCR4 positive cells enriches for definitive endoderm gene expression and depletes cells expressing markers of mesoderm, ectoderm and visceral endoderm

[0303] The $CXCR4^+$ and $CXCR4^-$ cells identified in Example 8 above were collected and analyzed for relative gene expression and the gene expression of the parent populations was determined simultaneously.

[0304] The relative levels of CXCR4 gene expression was dramatically increased with increasing dose of activin A (Figure 43). This correlated very well with the activn A dose-dependent increase of $CXCR4^+$ cells (Figures 42A-C). It is also clear that isolation of the $CXCR4^+$ cells from each population accounted for nearly all of the CXCR4 gene expression in that population. This demonstrates the efficiency of the FACS method for collecting these cells.

[0305] Gene expression analysis revealed that the $CXCR4^+$ cells contain not only the majority of the CXCR4 gene expression, but they also contained almost all of the gene expression for markers of definitive endoderm. As shown in Figures 44A-D, the $CXCR4^+$ cells were further enriched over the parent A100 population for SOX17, GSC, HNF3B, and MIXL1. In addition, the $CXCR4^-$ fraction contained very little gene expression for these definitive endoderm markers. Moreover, the $CXCR4^+$ and $CXCR4^-$ populations displayed the inverse pattern of gene expression for markers of mesoderm, ectoderm and extra-embryonic endoderm. Figures 45A-D shows that the $CXCR4^+$ cells were depleted for gene expression of Brachyury, MOX1, ZIC1, and SOX7 relative to the A100 parent population. This A100 parent population was already low in expression of these markers relative to the low dose or no activin A conditions. These results suggests that the isolation of $CXCR4^+$

cells from hESCs differentiated in the presence of high activin A yields a population that is highly enriched for and substantially pure definitive endoderm.

EXAMPLE 10

Quantitation of Definitive Endoderm Cells in a Cell Population Using CXCR4

[0306] To confirm the quantitation of the proportion of definitive endoderm cells present in a cell culture or cell population as determined previously herein and in United States Provisional Patent Application No. 60/532,004, entitled DEFINITIVE ENDODERM, filed December 23, 2003, the disclosure of which is incorporated herein by reference in its entirety, cells expressing CXCR4 and other markers of definitive endoderm were analyzed by FACS.

[0307] Using the methods described herein, hESCs were differentiated to produce definitive endoderm. Differentiated cultures were sorted by FACS using three cell surface epitopes, E-Cadherin, CXCR4, and Thrombomodulin. Sorted cell populations were then analyzed by Q-PCR to determine relative expression levels of markers for definitive and extraembryonic-endoderm as well as other cell types. CXCR4 sorted cells taken from optimally differentiated cultures resulted in the isolation of definitive endoderm cells that were >98% pure.

[0308] Table 2 shows the results of a marker analysis for a definitive endoderm culture that was differentiated from hESCs using the methods described herein.

Table 2

Composition of Definitive Endoderm Cultures

Marker(s)	Percent of culture	Percent Definitive Endoderm	Percent Extraembryonic endododerm	Percent hES cells
SOX17	70-80	100		
Thrombomodulin	<2	0	75	
AFP	<1	0	25	
CXCR4	70-80	100	0	
ECAD	10	0		100
other (ECAD neg.)	10-20			
Total	100	100	100	100

[0309] In particular, Table 2 indicates that CXCR4 and SOX17 positive cells (endoderm) comprised from 70%-80% of the cells in the cell culture. Of these SOX17-expressing cells, less than 2% expressed TM (pareital endoderm) and less than 1% expressed AFP (visceral endoderm). After subtracting the proportion of TM-positive and AFP-positive cells (combined parietal and visceral endoderm; 3% total) from the proportion of SOX17/CXCR4 positive cells, it can be seen that about 67% to about 77% of the cell culture was definitive endoderm. Approximately 10% of the cells were positive for E-Cadherin (ECAD), which is a marker for hESCs, and about 20% of the cells were of other cell types.

[0310] The methods, compositions, and devices described herein are presently representative of preferred embodiments and are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention and are defined by the scope of the disclosure. Accordingly, it will be apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

[0311] As used in the claims below and throughout this disclosure, by the phrase "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates

that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

References

[0312] Numerous literature and patent references have been cited in the present application. All references cited are incorporated by reference herein in their entireties.

[0313] For some references, the complete citation is in the body of the text. For other references the citation in the body of the text is by author and year, the complete citation being as follows:

[0314] Alexander, J., Rothenberg, M., Henry, G. L., and Stainier, D. Y. (1999). casanova plays an early and essential role in endoderm formation in zebrafish. Dev Biol *215*, 343-357.

[0315] Alexander, J., and Stainier, D. Y. (1999). A molecular pathway leading to endoderm formation in zebrafish. Curr Biol *9*, 1147-1157.

[0316] Aoki, T. O., Mathieu, J., Saint-Etienne, L., Rebagliati, M. R., Peyrieras, N., and Rosa, F. M. (2002). Regulation of nodal signalling and mesendoderm formation by TARAM-A, a TGFbeta-related type I receptor. Dev Biol *241*, 273-288.

[0317] Beck, S., Le Good, J. A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F., and Constam, D. B. (2002). Extra-embryonic proteases regulate Nodal signalling during gastrulation. Nat Cell Biol *4*, 981-985.

[0318] Beddington, R. S., Rashbass, P., and Wilson, V. (1992). Brachyury--a gene affecting mouse gastrulation and early organogenesis. Dev Suppl, 157-165.

[0319] Bongso, A., Fong, C. Y., Ng, S. C., and Ratnam, S. (1994). Isolation and culture of inner cell mass cells from human blastocysts. Hum Reprod *9*, 2110-2117.

[0320] Chang, H., Brown, C. W., and Matzuk, M. M. (2002). Genetic analysis of the mammalian transforming growth factor-beta superfamily. Endocr Rev *23*, 787-823.

[0321] Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., and Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. Development *120*, 1919-1928.

[0322] Dougan, S. T., Warga, R. M., Kane, D. A., Schier, A. F., and Talbot, W. S. (2003). The role of the zebrafish nodal-related genes squint and cyclops in patterning of mesendoderm. Development *130*, 1837-1851.

[0323] Feldman, B., Gates, M. A., Egan, E. S., Dougan, S. T., Rennebeck, G., Sirotkin, H. I., Schier, A. F., and Talbot, W. S. (1998). Zebrafish organizer development and germ-layer formation require nodal-related signals. Nature *395*, 181-185.

[0324] Futaki, S., Hayashi, Y., Yamashita, M., Yagi, K., Bono, H., Hayashizaki, Y., Okazaki, Y., and Sekiguchi, K. (2003). Molecular basis of constitutive production of basement membrane components: Gene expression profiles of engelbreth-holm-swarm tumor and F9 embryonal carcinoma cells. J Biol Chem.

[0325] Grapin-Botton, A., and Melton, D. A. (2000). Endoderm development: from patterning to organogenesis. Trends Genet *16*, 124-130.

[0326] Harris, T. M., and Childs, G. (2002). Global gene expression patterns during differentiation of F9 embryonal carcinoma cells into parietal endoderm. Funct Integr Genomics *2*, 105-119.

[0327] Hogan, B. L. (1996). Bone morphogenetic proteins in development. Curr Opin Genet Dev *6*, 432-438.

[0328] Hogan, B. L. (1997). Pluripotent embryonic cells and methods of making same (U.S.A., Vanderbilt University).

[0329] Howe, C. C., Overton, G. C., Sawicki, J., Solter, D., Stein, P., and Strickland, S. (1988). Expression of SPARC/osteonectin transcript in murine embryos and gonads. Differentiation *37*, 20-25.

[0330] Hudson, C., Clements, D., Friday, R. V., Stott, D., and Woodland, H. R. (1997). Xsox17alpha and -beta mediate endoderm formation in Xenopus. Cell *91*, 397-405.

[0331] Imada, M., Imada, S., Iwasaki, H., Kume, A., Yamaguchi, H., and Moore, E. E. (1987). Fetomodulin: marker surface protein of fetal development which is modulatable by cyclic AMP. Dev Biol *122*, 483-491.

[0332] Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P., and Hayashi, Y. (2002).

Depletion of definitive gut endoderm in Sox17-null mutant mice. Development *129*, 2367-2379.

[0333] Katoh, M. (2002). Expression of human SOX7 in normal tissues and tumors. Int J Mol Med *9*, 363-368.

[0334] Kikuchi, Y., Agathon, A., Alexander, J., Thisse, C., Waldron, S., Yelon, D., Thisse, B., and Stainier, D. Y. (2001). casanova encodes a novel Sox-related protein necessary and sufficient for early endoderm formation in zebrafish. Genes Dev *15*, 1493-1505.

[0335] Kimelman, D., and Griffin, K. J. (2000). Vertebrate mesendoderm induction and patterning. Curr Opin Genet Dev *10*, 350-356.

[0336] Kumar, A., Novoselov, V., Celeste, A. J., Wolfman, N. M., ten Dijke, P., and Kuehn, M. R. (2001). Nodal signaling uses activin and transforming growth factor-beta receptor-regulated Smads. J Biol Chem *276*, 656-661.

[0337] Labosky, P. A., Barlow, D. P., and Hogan, B. L. (1994a). Embryonic germ cell lines and their derivation from mouse primordial germ cells. Ciba Found Symp *182*, 157-168; discussion 168-178.

[0338] Labosky, P. A., Barlow, D. P., and Hogan, B. L. (1994b). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. Development *120*, 3197-3204.

[0339] Lickert, H., Kutsch, S., Kanzler, B., Tamai, Y., Taketo, M. M., and Kemler, R. (2002). Formation of multiple hearts in mice following deletion of beta-catenin in the embryonic endoderm. Dev Cell *3*, 171-181.

[0340] Lu, C. C., Brennan, J., and Robertson, E. J. (2001). From fertilization to gastrulation: axis formation in the mouse embryo. Curr Opin Genet Dev *11*, 384-392.

[0341] Miyazono, K., Kusanagi, K., and Inoue, H. (2001). Divergence and convergence of TGF-beta/BMP signaling. J Cell Physiol *187*, 265-276.

[0342] Niwa, H. (2001). Molecular mechanism to maintain stem cell renewal of ES cells. Cell Struct Funct *26*, 137-148.

[0343] Ogura, H., Aruga, J., and Mikoshiba, K. (2001). Behavioral abnormalities of Zic1 and Zic2 mutant mice: implications as models for human neurological disorders. Behav Genet *31*, 317-324.

[0344] Reubinoff, B. E., Pera, M. F., Fong, C. Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol *18*, 399-404.

[0345] Rodaway, A., and Patient, R. (2001). Mesendoderm. an ancient germ layer? Cell *105*, 169-172.

[0346] Rodaway, A., Takeda, H., Koshida, S., Broadbent, J., Price, B., Smith, J. C., Patient, R., and Holder, N. (1999). Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-beta family signals and discrimination of mesoderm and endoderm by FGF. Development *126*, 3067-3078.

[0347] Rohr, K. B., Schulte-Merker, S., and Tautz, D. (1999). Zebrafish zic1 expression in brain and somites is affected by BMP and hedgehog signalling. Mech Dev *85*, 147-159.

[0348] Schier, A. F. (2003). Nodal signaling in vertebrate development. Annu Rev Cell Dev Biol *19*, 589-621.

[0349] Schoenwolf, G. C., and Smith, J. L. (2000). Gastrulation and early mesodermal patterning in vertebrates. Methods Mol Biol *135*, 113-125.

[0350] Shamblott, M. J., Axelman, J., Wang, S., Bugg, E. M., Littlefield, J. W., Donovan, P. J., Blumenthal, P. D., Huggins, G. R., and Gearhart, J. D. (1998). Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc Natl Acad Sci U S A *95*, 13726-13731.

[0351] Shapiro, A. M., Lakey, J. R., Ryan, E. A., Korbutt, G. S., Toth, E., Warnock, G. L., Kneteman, N. M., and Rajotte, R. V. (2000). Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. N Engl J Med *343*, 230-238.

[0352] Shapiro, A. M., Ryan, E. A., and Lakey, J. R. (2001a). Pancreatic islet transplantation in the treatment of diabetes mellitus. Best Pract Res Clin Endocrinol Metab *15*, 241-264.

[0353] Shapiro, J., Ryan, E., Warnock, G. L., Kneteman, N. M., Lakey, J., Korbutt, G. S., and Rajotte, R. V. (2001b). Could fewer islet cells be transplanted in type 1 diabetes? Insulin independence should be dominant force in islet transplantation. Bmj *322*, 861.

[0354] Shiozawa, M., Hiraoka, Y., Komatsu, N., Ogawa, M., Sakai, Y., and Aiso, S. (1996). Cloning and characterization of Xenopus laevis xSox7 cDNA. Biochim Biophys Acta *1309*, 73-76.

[0355] Smith, J. (1997). Brachyury and the T-box genes. Curr Opin Genet Dev *7*, 474-480.

[0356] Smith, J. C., Armes, N. A., Conlon, F. L., Tada, M., Umbhauer, M., and Weston, K. M. (1997). Upstream and downstream from Brachyury, a gene required for vertebrate mesoderm formation. Cold Spring Harb Symp Quant Biol *62*, 337-346.

[0357] Takash, W., Canizares, J., Bonneaud, N., Poulat, F., Mattei, M. G., Jay, P., and Berta, P. (2001). SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. Nucleic Acids Res *29*, 4274-4283.

[0358] Taniguchi, K., Hiraoka, Y., Ogawa, M., Sakai, Y., Kido, S., and Aiso, S. (1999). Isolation and characterization of a mouse SRY-related cDNA, mSox7. Biochim Biophys Acta *1445*, 225-231.

[0359] Technau, U. (2001). Brachyury, the blastopore and the evolution of the mesoderm. Bioessays *23*, 788-794.

[0360] Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

[0361] Tremblay, K. D., Hoodless, P. A., Bikoff, E. K., and Robertson, E. J. (2000). Formation of the definitive endoderm in mouse is a Smad2-dependent process. Development *127*, 3079-3090.

[0362] Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-

PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, RESEARCH0034.

[0363] Varlet, I., Collignon, J., and Robertson, E. J. (1997). nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. Development *124*, 1033-1044.

[0364] Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P., and Robertson, E. J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. Genes Dev *17*, 1646-1662.

[0365] Weiler-Guettler, H., Aird, W. C., Rayburn, H., Husain, M., and Rosenberg, R. D. (1996). Developmentally regulated gene expression of thrombomodulin in postimplantation mouse embryos. Development *122*, 2271-2281.

[0366] Weiler-Guettler, H., Yu, K., Soff, G., Gudas, L. J., and Rosenberg, R. D. (1992). Thrombomodulin gene regulation by cAMP and retinoic acid in F9 embryonal carcinoma cells. Proceedings Of The National Academy Of Sciences Of The United States Of America *89*, 2155-2159.

[0367] Wells, J. M., and Melton, D. A. (1999). Vertebrate endoderm development. Annu Rev Cell Dev Biol *15*, 393-410.

[0368] Wells, J. M., and Melton, D. A. (2000). Early mouse endoderm is patterned by soluble factors from adjacent germ layers. Development *127*, 1563-1572.

[0369] Willison, K. (1990). The mouse Brachyury gene and mesoderm formation. Trends Genet *6*, 104-105.

[0370] Zhao, G. Q. (2003). Consequences of knocking out BMP signaling in the mouse. Genesis *35*, 43-56.

[0371] Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L., and Kuehn, M. R. (1993). Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. Nature *361*, 543-547.

[0372] D'Amour, et al. (2003) U.S. Provisional Patent Application No. 60/532,004

[0373] Kanai-Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PP, Hayashi Y. (2002) Depletion of definitive gut endoderm in SOX17-null mutant mice. Development. *129*,2367-79.

[0374] Kubo A, Shinozaki K, Shannon JM, Kouskoff V, Kennedy M, Woo S, Fehling HJ, Keller G. (2004) Development of definitive endoderm from embryonic stem cells in culture. Development. *131*,1651-62

[0375] McGrath KE, Koniski AD, Maltby KM, McGann JK, Palis J. (1999) Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4. Dev Biol. *213*, 442-56.

[0376] Nagasawa, T., Hirota, S., Tachibana, K., Takakura, N., Nishikawa, S., Kitamura, Y., Yoshida, N., Kikutani, H., and Kishimoto, T. (1996). Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. Nature *382*, 635-638.

[0377] Ma, Q., Jones, D., and Springer, T. A. (1999). The chemokine receptor CXCR4 is required for the retention of B lineage and granulocytic precursors within the bone marrow microenvironment. Immunity *10*, 463-471.

[0378] Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science *272*, 872-877.

[0379] Kim, C. H., and Broxmeyer, H. E. (1999). Chemokines: signal lamps for trafficking of T and B cells for development and effector function. J Leukoc Biol *65*, 6-15.

WHAT IS CLAIMED IS:

1. A mammalian cell culture comprising endodermal cells, wherein the expression of a marker selected from the group consisting of CXCR4 and SOX17 is greater than the expression of the alpha-fetoprotein (AFP) and the Thrombomodulin (TM) markers in at least about 90% of said endodermal cells.

CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM

Abstract of the Disclosure

Disclosed herein are compositions comprising definitive endoderm cells and methods of producing the same. Also disclosed herein are compositions comprising substantially purified definitive endoderm cells as well as methods for enriching, isolating and purifying definitive endoderm cells from other cell types.

Figure 1

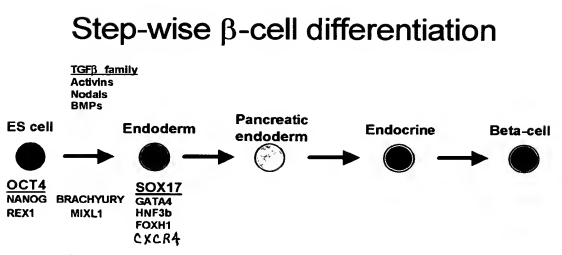

Step-wise β-cell differentiation
TGFβ family
Activins
Nodals
BMPs
ES cell
Endoderm
Pancreatic endoderm
Endocrine
Beta-cell
OCT4
NANOG
REX1
BRACHYURY
MIXL1
SOX17
GATA4
HNF3b
FOXH1
CXCR4

Figure 2

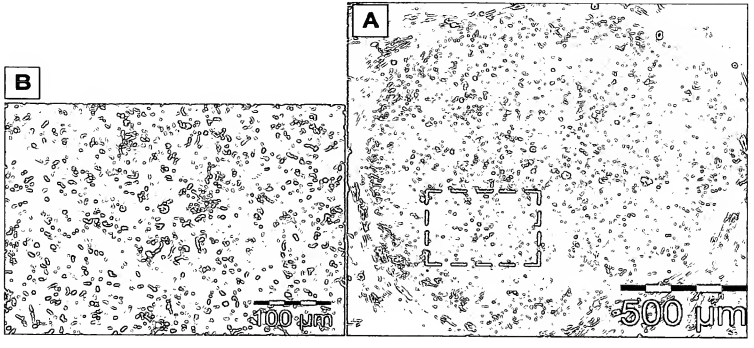
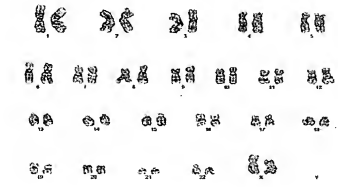


Figure 3

hESCyT-25 Chromosome Analysis at Passage 10



hESCyT-25 Chromosome Analysis at Passage 28

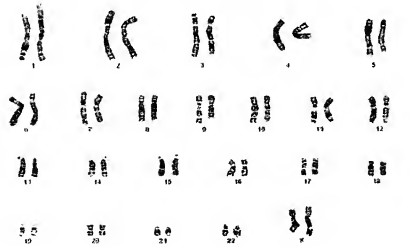


Figure 4

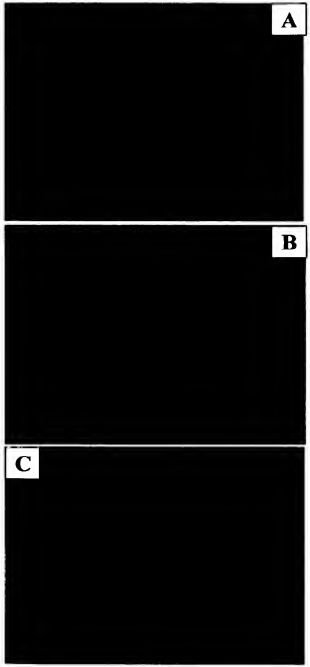


Figure 5

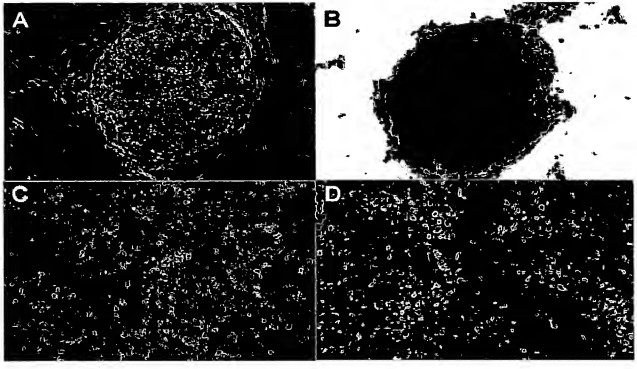


Figure 6

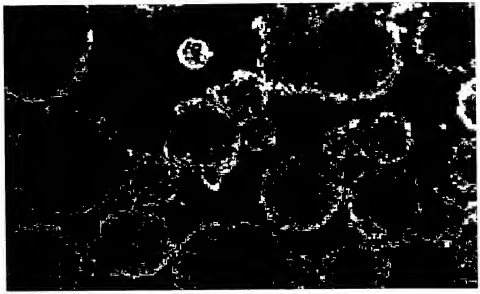


Figure 7

Figure 8

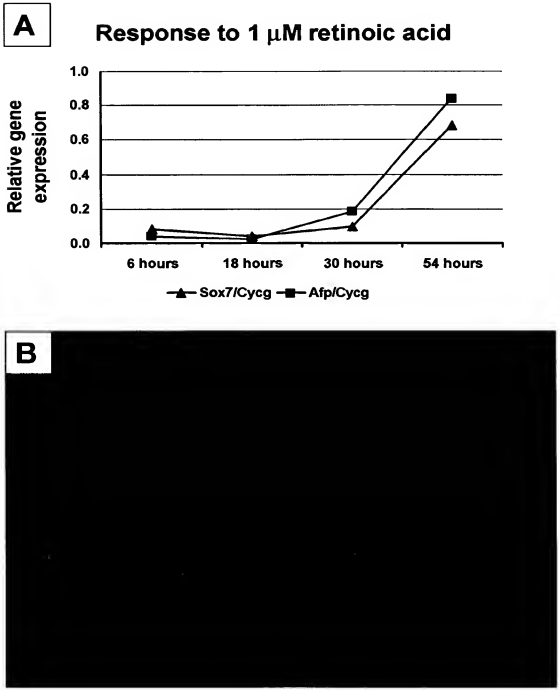

A
Response to 1 μM retinoic acid
Relative gene expression
1.0
0.8
0.6
0.4
0.2
0.0
6 hours
18 hours
30 hours
54 hours
Sox7/Cycg
Afp/Cycg
B

Figure 9

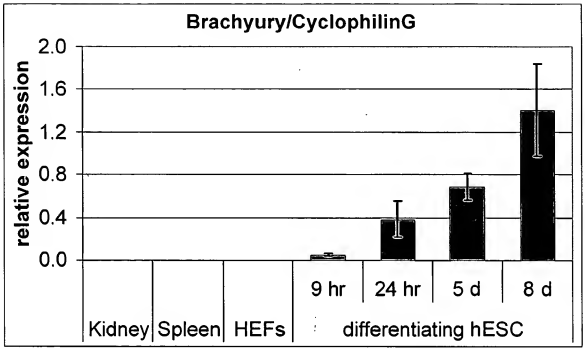
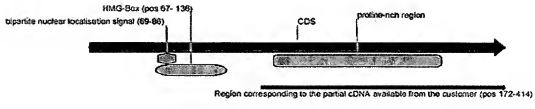
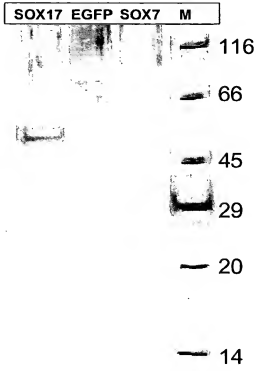

Brachyury/CyclophilinG
relative expression
2.0
1.6
1.2
0.8
0.4
0.0
9 hr
24 hr
5 d
8 d
Kidney
Spleen
HEFs
differentiating hESC

Figure 10



Translation of Sox17, hsa
414 aa

Figure 11



Anti- Sox17

Figure 12

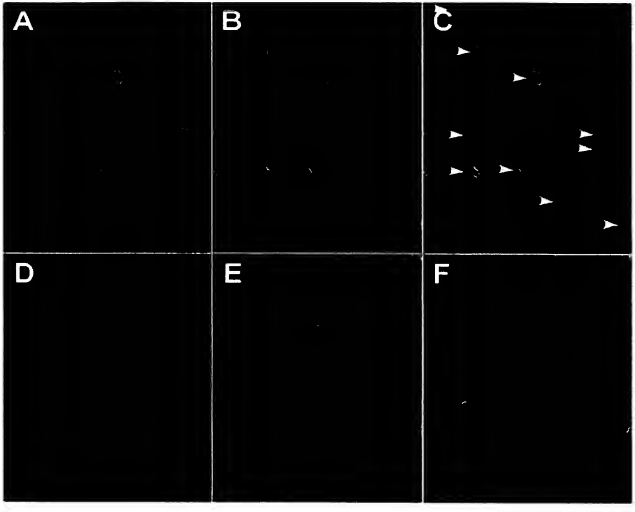


Figure 13

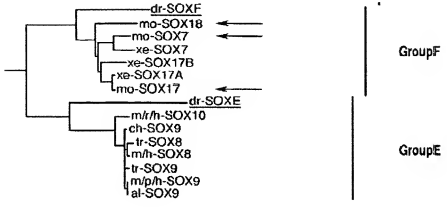
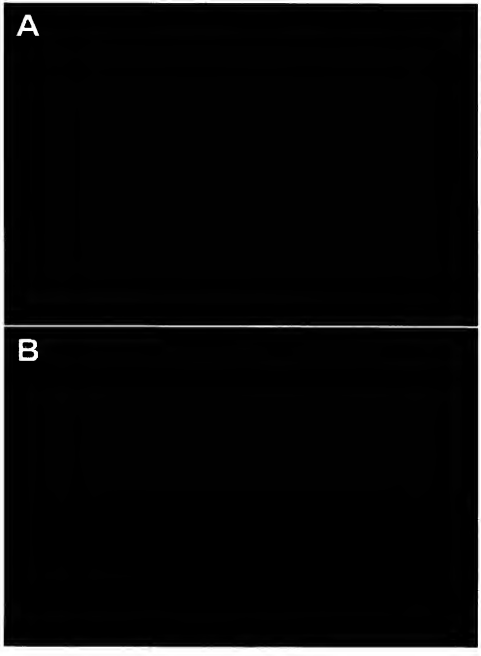

dr-SOXF
mo-SOX18
mo-SOX7
xe-SOX7
xe-SOX17B
xe-SOX17A
mo-SOX17
dr-SOXE
m/r/h-SOX10
ch-SOX9
tr-SOX8
m/h-SOX8
tr-SOX9
m/p/h-SOX9
al-SOX9
GroupF
GroupE

Figure 14



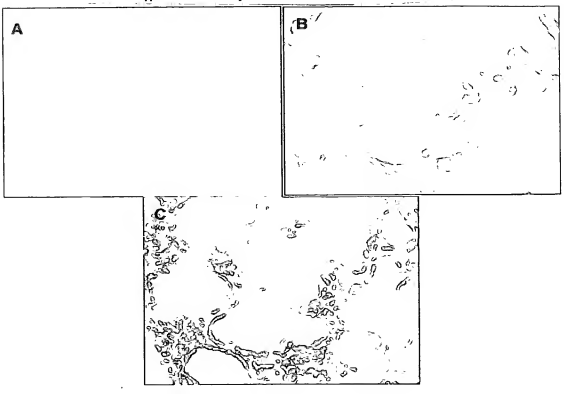


Figure 15

Figure 16

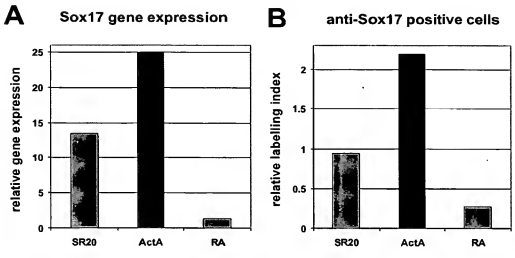

A
Sox17 gene expression
relative gene expression
25
20
15
10
5
0
SR20
ActA
RA
B
anti-Sox17 positive cells
relative labelling index
2
1.5
1
0.5
0
SR20
ActA
RA

Figure 17

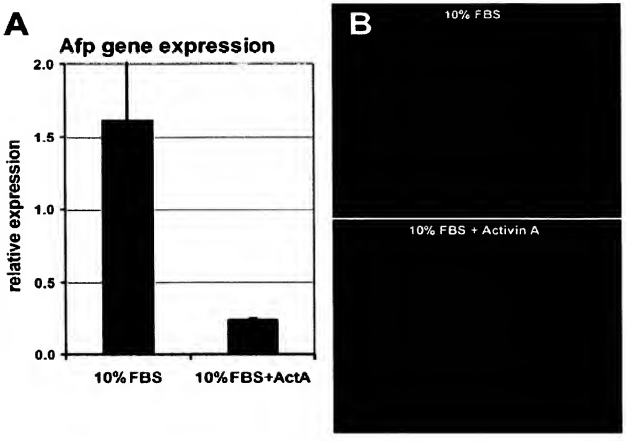

A
Afp gene expression
relative expression
2.0
1.5
1.0
0.5
0.0
10% FBS
10%FBS+ActA
B
10% FBS
10% FBS + Activin A

Figure 18

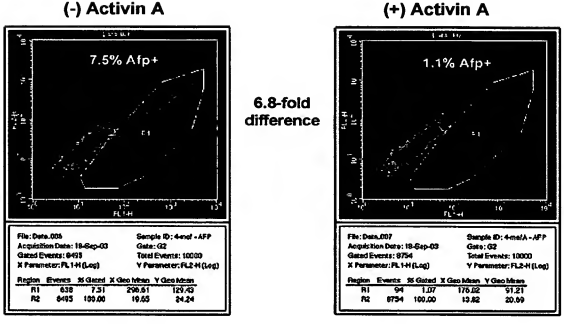


Figure 19

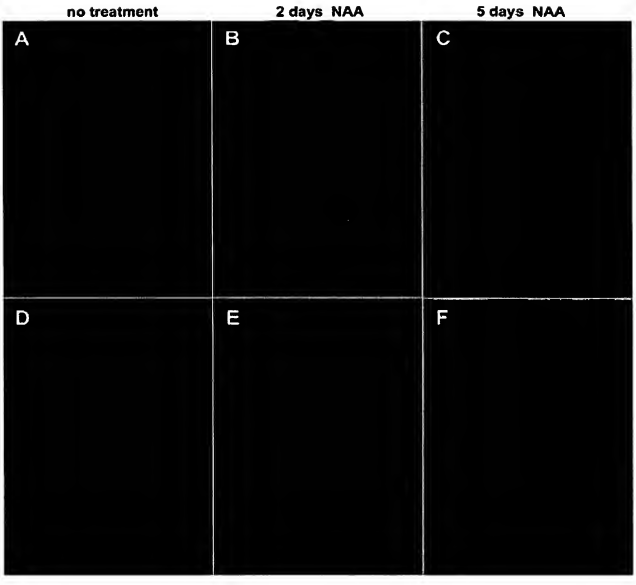


Figure 20

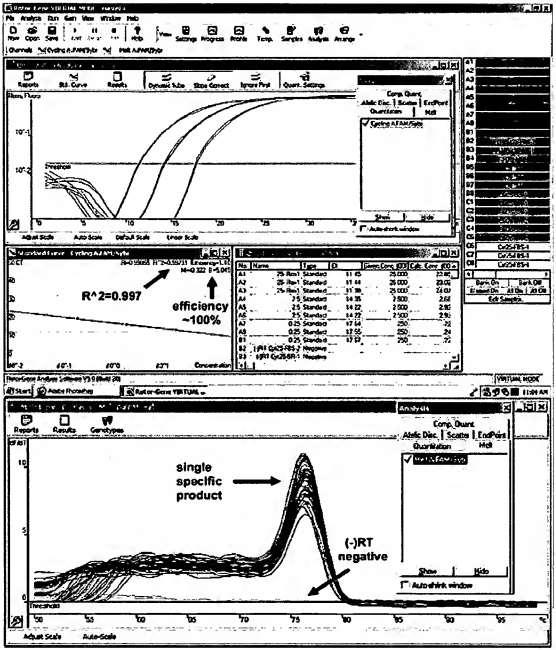

R^2=0.997
efficiency ~100%
single specific product
(-)RT negative

Figure 21

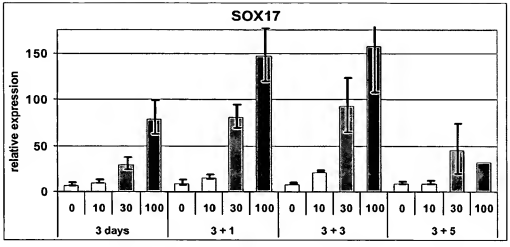

SOX17
relative expression
150
100
50
0
0
10
30
100
3 days
3 + 1
3 + 3
3 + 5

Figure 22A

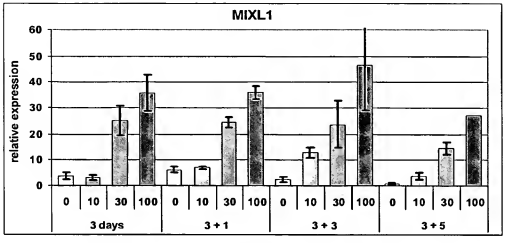


Figure 22B

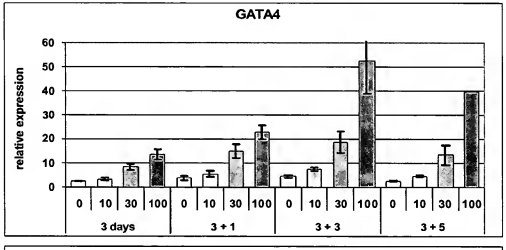


Figure 22C

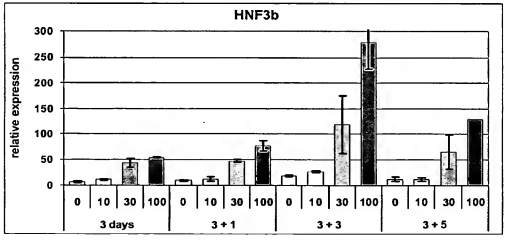


Figure 23A

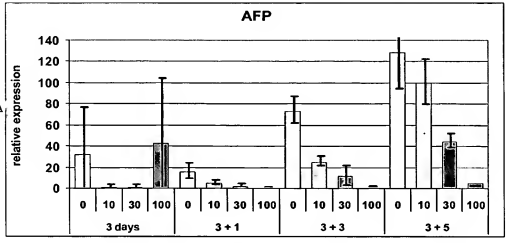


Figure 23B

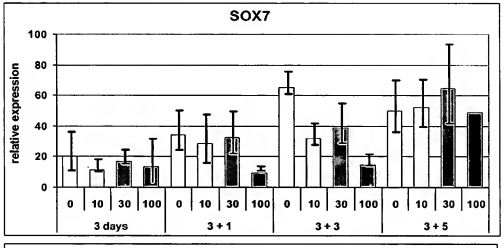


Figure 23C

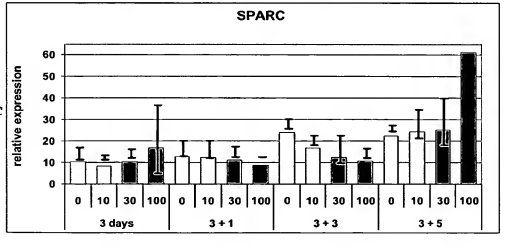


Figure 24A

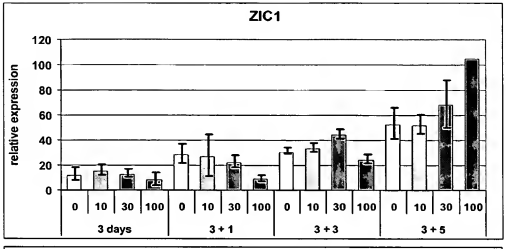


Figure 24B

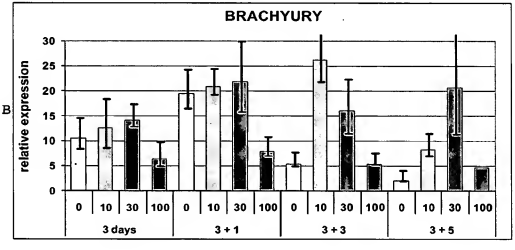


Figure 25

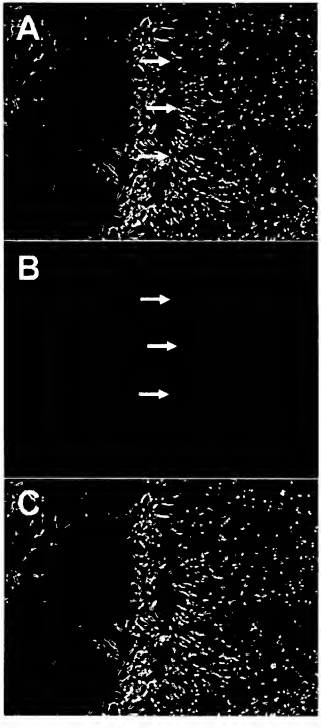


Figure 26

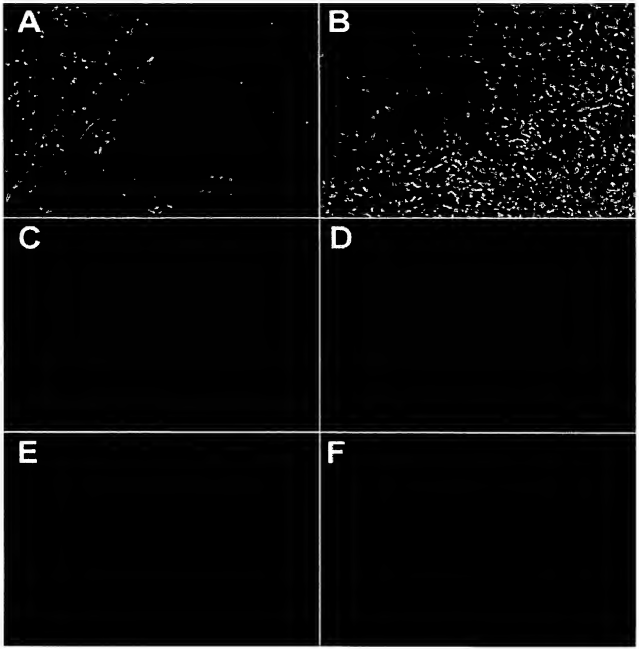


Figure 27

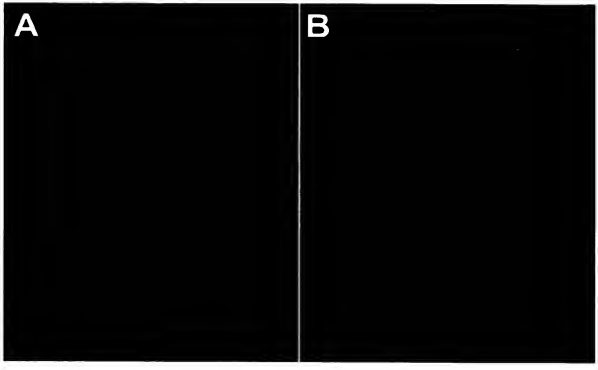


Figure 28

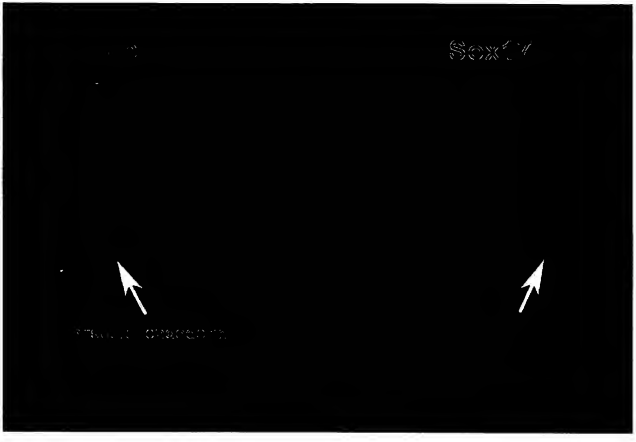


Figure 29

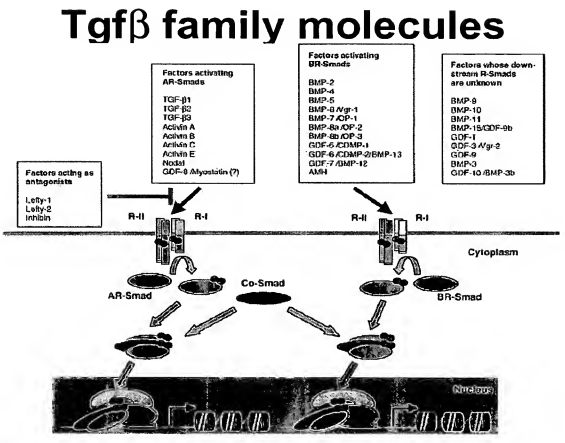

Tgfβ family molecules
Factors activating AR-Smads
TGF-β1
TGF-β2
TGF-β3
Activin A
Activin B
Activin C
Activin E
Nodal
GDF-8 /Myostatin (?)
Factors activating BR-Smads
BMP-2
BMP-4
BMP-5
BMP-6 /Vgr-1
BMP-7 /OP-1
BMP-8a /OP-2
BMP-8b /OP-3
GDF-5 /CDMP-1
GDF-6 /CDMP-2/BMP-13
GDF-7 /DMP-1Z
AMH
Factors whose down-stream R-Smads are unknown
BMP-9
BMP-10
BMP-11
BMP-15/GDF-9b
GDF-1
GDF-3 /Vgr-2
GDF-9
BMP-3
GDF-10 /BMP-3b
Factors acting as antagonists
Lefty-1
Lefty-2
Inhibin
R-II
R-I
R-II
R-I
Cytoplasm
AR-Smad
Co-Smad
BR-Smad
Nucleus

Figure 30

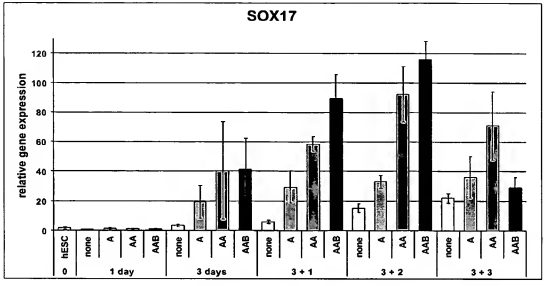

SOX17
relative gene expression
120
100
80
60
40
20
0
hESC
none
A
AA
AAB
0
1 day
3 days
3 + 1
3 + 2
3 + 3

Figure 31

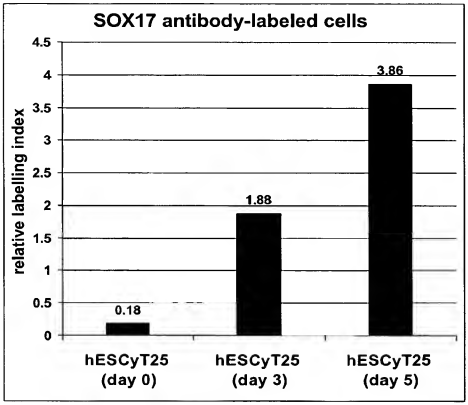

SOX17 antibody-labeled cells
relative labelling index
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0.18
1.88
3.86
hESCyT25 (day 0)
hESCyT25 (day 3)
hESCyT25 (day 5)

Figure 32

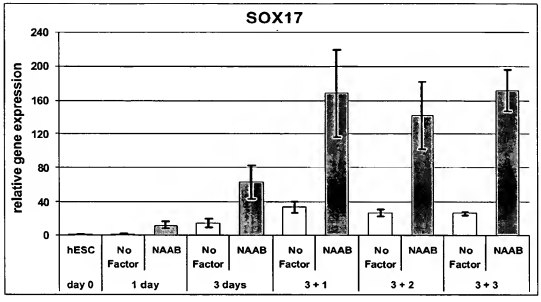

SOX17
relative gene expression
240
200
160
120
80
40
0
hESC
No Factor
NAAB
No Factor
NAAB
No Factor
NAAB
No Factor
NAAB
No Factor
NAAB
day 0
1 day
3 days
3 + 1
3 + 2
3 + 3

Figure 33

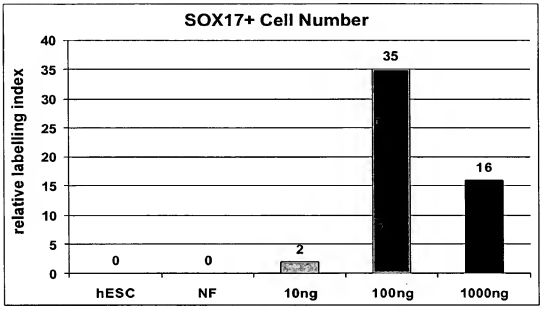

SOX17+ Cell Number
relative labelling index
40
35
30
25
20
15
10
5
0
0
0
2
35
16
hESC
NF
10ng
100ng
1000ng

Figure 34

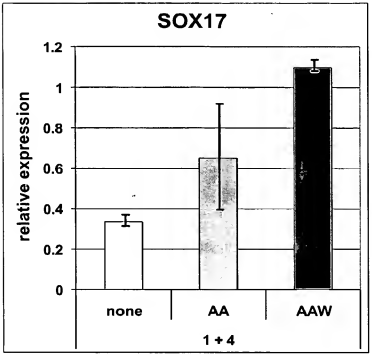


Figure 35A

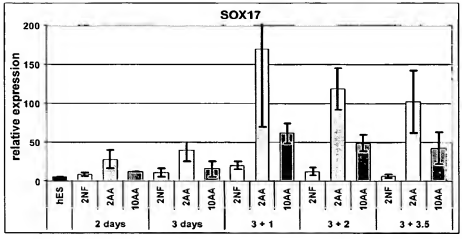


Figure 35B

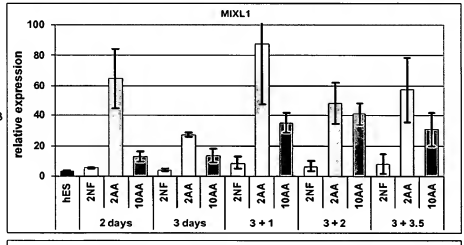


Figure 35C

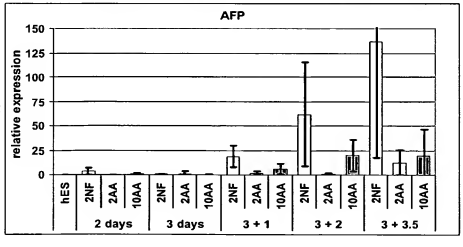


Figure 36

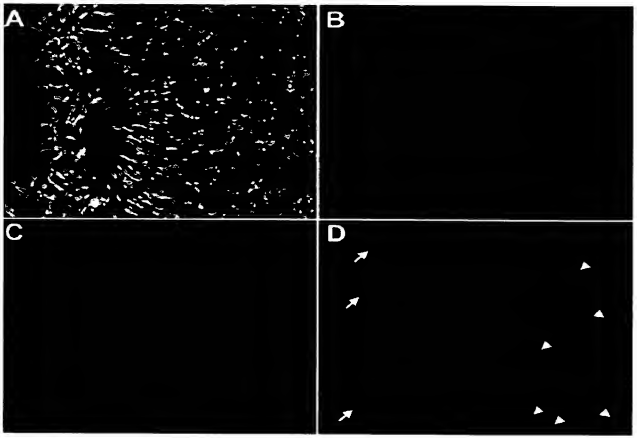


Figure 37

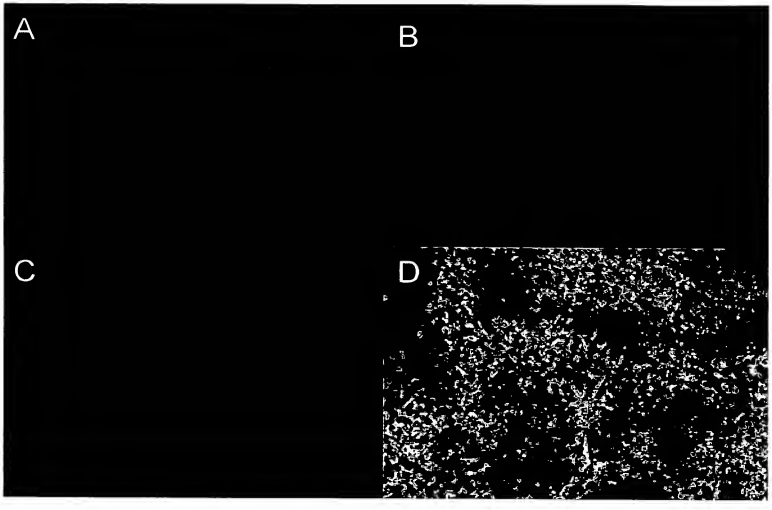


Figure 38

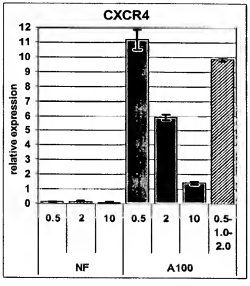

CXCR4
relative expression
12
11
10
9
8
7
6
5
4
3
2
1
0
0.5
2
10
0.5
2
10
0.5-
1.0-
2.0
NF
A100

Figure 39A

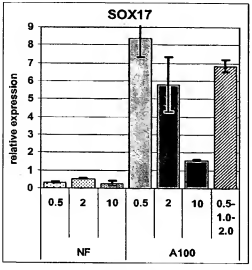
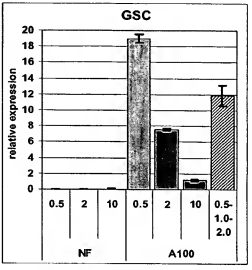


Figure 39B



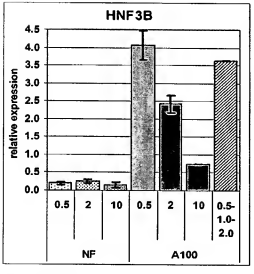


Figure 39C

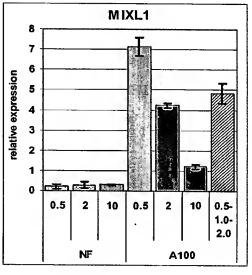


Figure 39D

Figure 40A

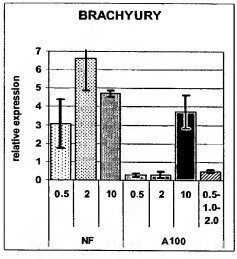


Figure 40B

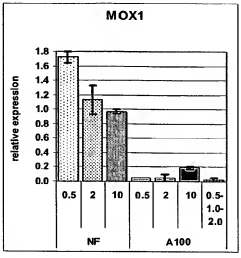


Figure 40C

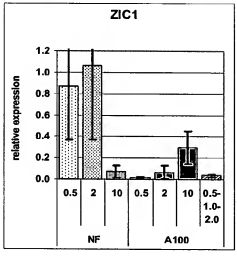


Figure 40D

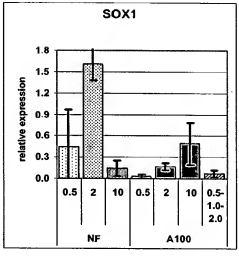


Figure 40E

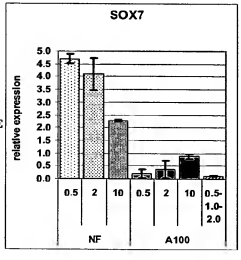


Figure 41

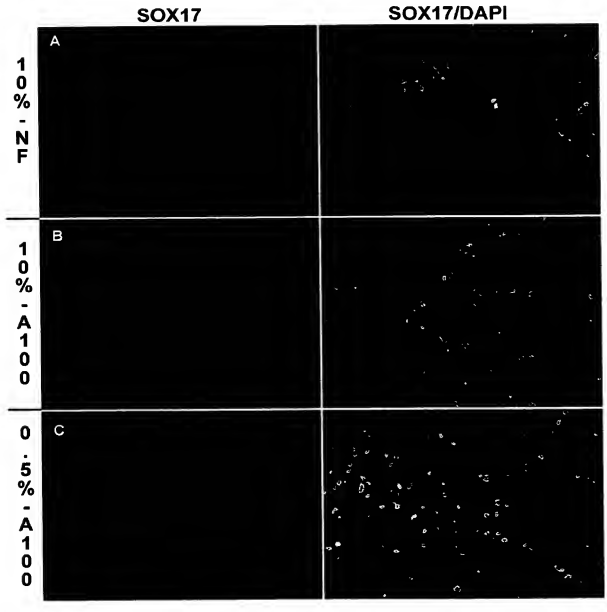


Figure 42

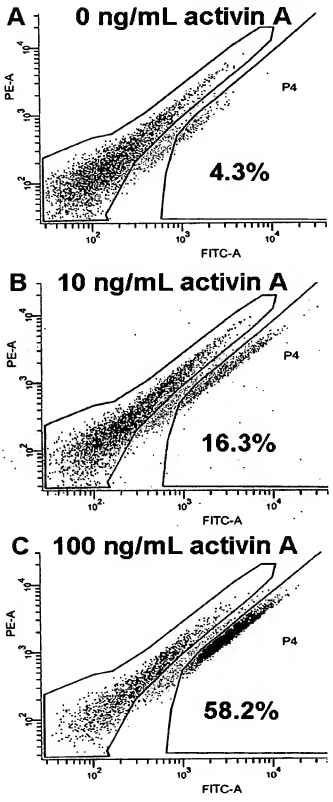

A
0 ng/mL activin A
PE-A
FITC-A
P4
4.3%
B
10 ng/mL activin A
PE-A
FITC-A
P4
16.3%
C
100 ng/mL activin A
PE-A
FITC-A
P4
58.2%

Figure 43

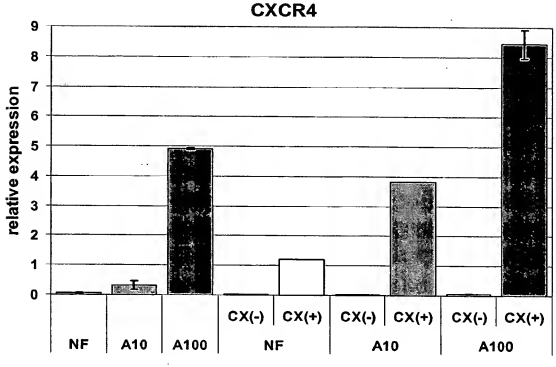

CXCR4
relative expression
9
8
7
6
5
4
3
2
1
0
CX(-)
CX(+)
CX(-)
CX(+)
CX(-)
CX(+)
NF
A10
A100
NF
A10
A100

Figure 44 A

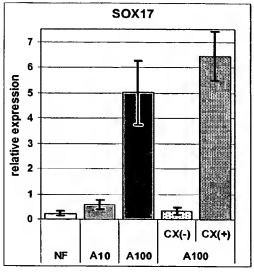
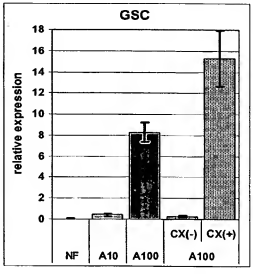


Figure 44B



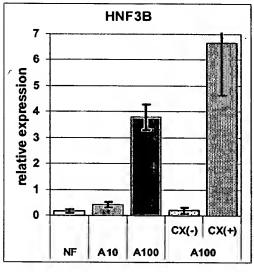


Figure 44C

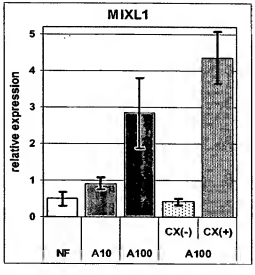


Figure 44D

Figure 45 A

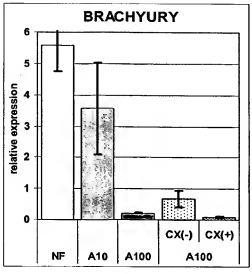
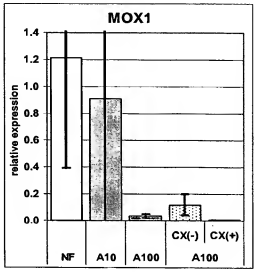


Figure 45B



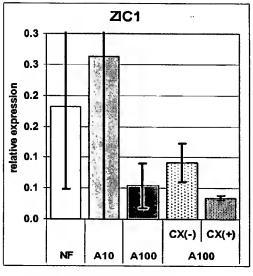


Figure 45C

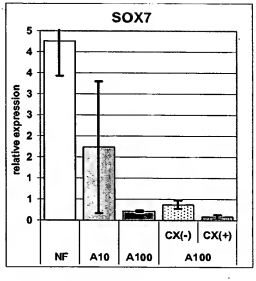


Figure 45D

SEQUENCE LISTING

<110> Emmanuel E. Baetge
Kevin Allen D'Amour
Alan D. Agulnick

<120> CHEMOKINE CELL SURFACE RECEPTOR FOR THE ISOLATION OF DEFINITIVE ENDODERM

<130> CYTHERA.045PR2

<160> 2

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1245
<212> DNA
<213> Homo sapiens

<400> 1
atgagcagcc cggatgcggg atacgccagt gacgaccaga gccagaccca gagcgcgctg 60
cccgcggtga tggccgggct gggcccctgc ccctgggccg agtcgctgag ccccatcggg 120
gacatgaagg tgaagggcga ggcgccggcg aacagcggag caccggccgg ggccgcgggc 180
cgagccaagg gcgagtcccg tatccggcgg ccgatgaacg ctttcatggt gtgggctaag 240
gacgagcgca agcggctggc gcagcagaat ccagacctgc acaacgccga gttgagcaag 300
atgctgggca agtcgtggaa ggcgctgacg ctggcggaga agcggccctt cgtggaggag 360
gcagagcggc tgcgcgtgca gcacatgcag gaccacccca actacaagta ccggccgcgg 420
cggcgcaagc aggtgaagcg gctgaagcgg gtggagggcg gcttcctgca cggcctggct 480
gagccgcagg cggccgcgct gggccccgag ggcggccgcg tggccatgga cggcctgggc 540
ctccagttcc ccgagcaggg cttccccgcc ggcccgccgc tgctgcctcc gcacatgggc 600
ggccactacc gcgactgcca gagtctgggc gcgcctccgc tcgacggcta cccgttgccc 660
acgcccgaca cgtccccgct ggacggcgtg gaccccgacc cggctttctt cgccgccccg 720
atgcccgggg actgcccggc ggccggcacc tacagctacg cgcaggtctc ggactacgct 780
ggcccccgg agcctcccgc cggtcccatg cacccccgac tcggcccaga gcccgcgggt 840
ccctcgattc cgggcctcct ggcgccaccc agcgcccttc acgtgtacta cggcgcgatg 900
ggctcgcccg gggcgggcgg cgggcgcggc ttccagatgc agccgcaaca ccagcaccag 960
caccagcacc agcaccaccc cccgggcccc ggacagccgt cgccccctcc ggaggcactg 1020
ccctgccggg acggcacgga ccccagtcag cccgccgagc tcctcgggga ggtggaccgc 1080
acggaatttg aacagtatct gcacttcgtg tgcaagcctg agatgggcct cccctaccag 1140
gggcatgact ccggtgtgaa tctccccgac agccacgggg ccatttcctc ggtggtgtcc 1200
gacgccagct ccgcggtata ttactgcaac tatcctgacg tgtga 1245

<210> 2
<211> 414
<212> PRT
<213> Homo sapiens

<400> 2
Met Ser Ser Pro Asp Ala Gly Tyr Ala Ser Asp Asp Gln Ser Gln Thr
1 5 10 15
Gln Ser Ala Leu Pro Ala Val Met Ala Gly Leu Gly Pro Cys Pro Trp
20 25 30
Ala Glu Ser Leu Ser Pro Ile Gly Asp Met Lys Val Lys Gly Glu Ala
35 40 45

```
Pro Ala Asn Ser Gly Ala Pro Ala Gly Ala Ala Gly Arg Ala Lys Gly
    50                  55                  60
Glu Ser Arg Ile Arg Arg Pro Met Asn Ala Phe Met Val Trp Ala Lys
65                  70                  75                  80
Asp Glu Arg Lys Arg Leu Ala Gln Gln Asn Pro Asp Leu His Asn Ala
                85                  90                  95
Glu Leu Ser Lys Met Leu Gly Lys Ser Trp Lys Ala Leu Thr Leu Ala
            100                 105                 110
Glu Lys Arg Pro Phe Val Glu Glu Ala Glu Arg Leu Arg Val Gln His
        115                 120                 125
Met Gln Asp His Pro Asn Tyr Lys Tyr Arg Pro Arg Arg Arg Lys Gln
    130                 135                 140
Val Lys Arg Leu Lys Arg Val Glu Gly Gly Phe Leu His Gly Leu Ala
145                 150                 155                 160
Glu Pro Gln Ala Ala Ala Leu Gly Pro Glu Gly Gly Arg Val Ala Met
                165                 170                 175
Asp Gly Leu Gly Leu Gln Phe Pro Glu Gln Gly Phe Pro Ala Gly Pro
            180                 185                 190
Pro Leu Leu Pro Pro His Met Gly Gly His Tyr Arg Asp Cys Gln Ser
        195                 200                 205
Leu Gly Ala Pro Pro Leu Asp Gly Tyr Pro Leu Pro Thr Pro Asp Thr
    210                 215                 220
Ser Pro Leu Asp Gly Val Asp Pro Asp Pro Ala Phe Phe Ala Ala Pro
225                 230                 235                 240
Met Pro Gly Asp Cys Pro Ala Ala Gly Thr Tyr Ser Tyr Ala Gln Val
                245                 250                 255
Ser Asp Tyr Ala Gly Pro Pro Glu Pro Pro Ala Gly Pro Met His Pro
            260                 265                 270
Arg Leu Gly Pro Glu Pro Ala Gly Pro Ser Ile Pro Gly Leu Leu Ala
        275                 280                 285
Pro Pro Ser Ala Leu His Val Tyr Tyr Gly Ala Met Gly Ser Pro Gly
    290                 295                 300
Ala Gly Gly Gly Arg Gly Phe Gln Met Gln Pro Gln His Gln His Gln
305                 310                 315                 320
His Gln His Gln His His Pro Pro Gly Pro Gly Gln Pro Ser Pro Pro
                325                 330                 335
Pro Glu Ala Leu Pro Cys Arg Asp Gly Thr Asp Pro Ser Gln Pro Ala
            340                 345                 350
Glu Leu Leu Gly Glu Val Asp Arg Thr Glu Phe Glu Gln Tyr Leu His
        355                 360                 365
Phe Val Cys Lys Pro Glu Met Gly Leu Pro Tyr Gln Gly His Asp Ser
    370                 375                 380
Gly Val Asn Leu Pro Asp Ser His Gly Ala Ile Ser Ser Val Val Ser
385                 390                 395                 400
Asp Ala Ser Ser Ala Val Tyr Tyr Cys Asn Tyr Pro Asp Val
                405                 410
```